IMPLEMENTATION AND SIMULATION OF MOBILE SENSOR NETWORKS FOR NUCLEAR RADIATION DETECTION

BY

JIFU ZHAO

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Nuclear, Plasma, and Radiological Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Rizwan Uddin, Chair
Assistant Professor Shiva Abbaszadeh
Professor Robert J. Brunner
Assistant Professor Kathryn D. Huff
Adjunct Research Assistant Professor Clair J. Sullivan

Abstract

From preventing the threat of nuclear weapons proliferation to monitoring the transportation of special nuclear materials, nuclear radiation detection plays an important role in national security applications. However, changing background radiation, shielding effects, and short collection time make radiation detection a challenging problem. Anomaly detection, source localization, and isotope identification are three major parts of radiation detection. The concept of mobile radiation sensor networks, which utilize multiple mobile radiation detectors, has been proposed to solve these problems. This work mainly focuses on developing and testing methodologies for anomaly detection and radioactive source localization using mobile sensor networks. A collection of techniques and analyses for radiation detection are presented and evaluated. More specifically, in this work, a mobile sensor network simulation system is first developed to simulate the scenario where multiple radiation detectors move around a city. Based on the simulated data, the performance characteristics of mobile sensor networks for radiation detection are studied and quantified. Next, focusing on geospatial modeling of radiation count data, Poisson kriging is proposed to estimate the background radiation level and perform anomalous source detection. The proposed method is validated using simulated source data injected in measured background radiation data and results indicate that the proposed algorithm can detect the anomalous radiation source with 90% accuracy under certain conditions. Additionally, source localization techniques based on maximum likelihood estimation are explored in detail. Simulation and experimental results show that source localization error can be reduced to be within 3 meters. Lastly, an exploratory study of spectrum-based anomaly detection techniques is presented. The performance of different machine learning techniques is evaluated and compared using simulated radiation data.

To My Family.

Acknowledgments

First, I would like to express my sincere thanks to my advisor, Professor Rizwan Uddin. This work would not have been possible without his support and guidance. Also, I want to express my thanks to Professor Clair Sullivan for her guidance and help during the first three years.

I would like to express my thanks to my committee members, Professor Shiva Abbaszadeh, Professor Robert Brunner, and Professor Kathryn Huff, for investing time and efforts to help me improve this thesis.

I want to thank my group members and alumni for their instructive discussions and for their kind help in the past five years.

Last but not least, I would like to thank my parents and Chunyan Hao for their endless love and support on my road to knowledge.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ALSO	Attribute-wise Learning for Scoring Outliers
AUC	Area Under the ROC Curve or Area Under the Curve
CNN	Convolutional Neural Network
D3S	Discreet Dual Detector
DSN	Distributed Sensor Networks
FPR	False Positive Rate
GADRAS	Gamma Detector Response and Analysis Software
IAEA	International Atomic Energy Agency
ITDB	Incident and Trafficking Database
ITP	Iterative Pruning
KDE	Kernel Density Estimation
KNN	K-nearest Neighbor
LOF	Local Outlier Factor
MAE	Mean Absolute Error
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimation
MoE	Mean of Estimators
MoM	Mean of Measurements
MSE	Mean Squared Error
NORM	Naturally Occurring Radioactive Material
PCA	Principal Component Analysis
ReLU	Rectified Linear Unit
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
SNIP	Statistics-sensitive Non-linear Iterative Peak-clipping

SNM	Special Nuclear Material
SNR	Signal-to-Noise Ratio
SPRT	Sequential Probability Ratio Test
SVD	Singular Value Decomposition
SVM	Support Vector Machine
TPR	True Positive Rate

Chapter 1

Introduction

1.1 Nuclear Radiation Detection Problems

Nuclear weapons, radiological dispersal devices, as well as lost radioactive sources are threats to national security and human health. According to the IAEA Incident and Trafficking Database (ITDB)[1], from 1993 to 2017, there were 3,235 confirmed incidents reported in total, including incidents related to trafficking or malicious use. There is a need to develop nuclear radiation detection and monitoring systems and algorithms that can prevent potential dangers. Currently, large radiation detection systems have been deployed at key choke points such as airports and customs to prevent the unauthorized entry or exit of radioactive materials. However, these systems at the borders cannot help to monitor materials already in the country, such as radioactive sources for medical, industrial, or research purposes. To enable nuclear radiation detection and monitoring over large areas, such as big cities, a distributed radiation detection system is needed.

Different from other signal capturing processes, there are some unique challenges in nuclear radiation detection. One of the key challenges is background radiation. Background radiation comes from the environment and it exists almost everywhere. It is influenced by many factors, such as weather conditions and surrounding buildings. Background radiation can be seen as the noise, which keeps changing temporally and spatially. The signal that radiation monitoring agencies are mostly interested in comes from special radioactive materials such as ^{235}U and Pu. Due to the inverse square law of radiation propagation [1], the signal-to-noise ratio (SNR) decreases dramatically as the distance from radiation detectors to radioactive sources increases, which makes it hard to collect useful signals at a large distance. Another challenge is the shielding effect. Different materials, such as buildings and automobiles, can reduce the signal strength accordingly and make it even harder to detect. In addition, limited data collection time can be a challenge. Radioactive materials may be carried by moving vehicles or people. This requires the data collection time to be as short as few to several seconds. All these factors make radiation detection and monitoring a very challenging problem.

[1]International Atomic Energy Agency (IAEA) Incident and Trafficking Database (ITDB) 2018 Fact Sheet: `https://www.iaea.org/sites/default/files/18/12/itdb-factsheet-2018.pdf`.

A typical radiation detection system can be divided into several parts. A schematic diagram is shown in Fig. 1.1. Radiation signals that might contain information from radioactive sources are first collected by radiation detectors, which is denoted by "data collection" process. Once collected, the data will be analyzed by various algorithms and techniques. The process of radiation detection can be divided into three distinct parts: anomaly detection, source localization, and isotope identification. The goal of anomaly detection is to distinguish the signals that contain radioactive source information from the set of measured radiation signals. After anomaly detection, abnormal measurements are separated and the data that need to be further processed will be dramatically decreased in volume since most data contain only background radiation information. Also, smaller regions that may contain radioactive sources are identified. The goal of source localization is to accurately locate the radioactive sources based on the information from anomaly detection. The last one, isotope identification, is to identify the isotopic composition of the radioactive source. For nuclear radiation, data collection is carried out by radiation detectors. High quality data can help simplify subsequent steps. However, this work is not aimed at high quality radiation detector design or analysis. Instead, this thesis focuses on the analysis of radiation signals coming from large number of low-quality detectors. More specifically, the anomaly detection and source localization steps are analyzed in detail. Isotope identification, on the other hand, is not included in this thesis.

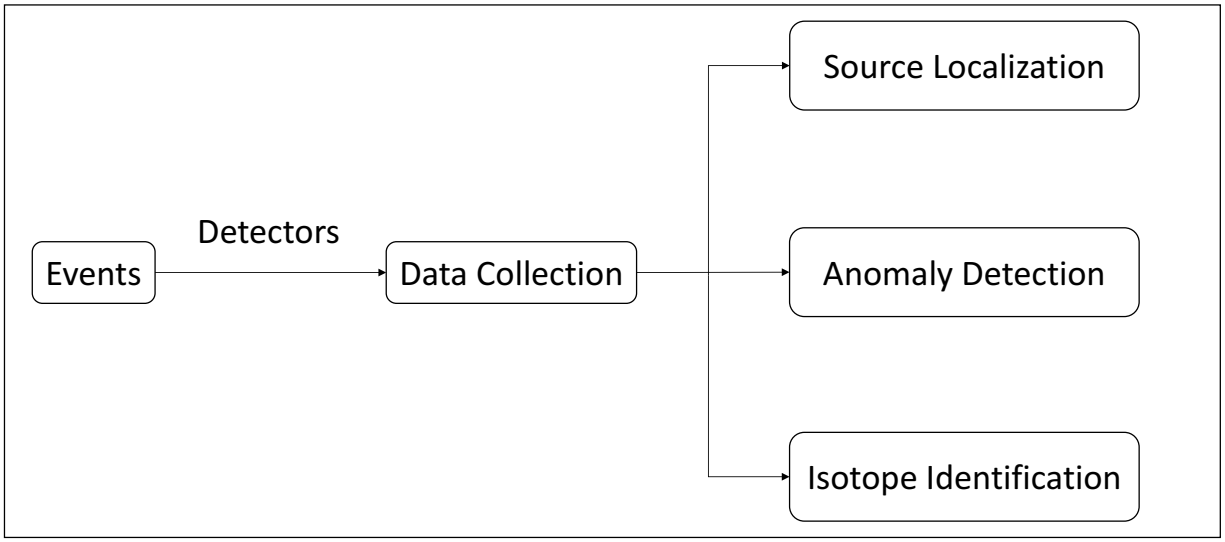


Figure 1.1: Schematic diagram of typical radiation detection system.

1.2 Mobile Sensor Networks

To enable nuclear radiation detection over large areas, a single or even several radiation detectors are insufficient. The concept of distributed sensor networks (DSNs) [2] that utilize multiple stationary detectors has been put forward and well studied. Distributed sensor networks have some advantages over single detector systems. Nemzed et al. [3] noticed that when combining the data from a small number of detectors, the signal-to-noise ratio (SNR) increased as $\sqrt{N}$, where N is the number of detectors used. Brennan et al. have shown, through validated simulation, that distributed detector systems improved the performance over a single detector [2], even when the detectors' sensitivity or energy resolution were not as good as the single detector [4]. Through mathematical analysis and simulations, Chandy et al. found that the networks of static sensors can reduce the rate of false positives and help direct people to rapidly locate the radioactive sources [5].

However, due to the inverse square law of radiation propagation, SNR decreases dramatically as the distance from radiation detectors to the actual source increases. This requires the distributed sensor network system to have a high density of sensors to ensure high detectability of radioactive sources. In other words, with the same number of detectors, to ensure sufficient detectability, the area that can be covered by the distributed sensor network is restricted, which will increase the cost when the area to be monitored is large. To solve these problems, there has been research focused on using mobile sensor networks to detect radioactive sources [6–9].

Mobile detectors are easily deployable. Small mobile detectors can be carried by people and larger detectors can be carried by vehicles such as police cars. In addition, with the same number of detectors, due to the mobility, a mobile sensor network system can cover a larger area. Study has shown that mobile sensor networks are more effective in reducing the time to locate the radioactive sources than stationary sensor networks [6].

Mobility also leads to some challenges. Traditional techniques usually combine the information from surrounding detectors or use the historical information from fixed positions [10, 11] to make more robust estimations. However, with mobile sensor networks, the position of each detector changes over time and the historical data from fixed positions or the information from surrounding detectors may not be available. These factors require new algorithms for mobile sensor networks, which is the major focus of this work.

1.3 Current Radiation Detection Methods

Traditionally, there are two kinds of radiation detection: active and passive [12]. In active detection, externally generated bremsstrahlung X-rays [13] or neutrons [14] are needed to investigate the suspicious materials. Special nuclear materials (SNM) can be identified upon analyzing the spectra [13, 14]. Active detection is generally used at airports and customs, but it is expensive and lacks mobility, which make it unsuitable for large-scale nuclear radiation detection. On the other hand, passive detection does not need neutrons or photons to probe, and it measures the emitted radiation directly. Passive mode is widely used in radiation detection. The existing algorithms, which either analyze the gross count rate or the spectra, are mainly based on passive detection. This thesis only explores passive detection.

Existing radiation detection methods mainly focus on three aspects: anomaly detection, isotope identification, and source localization. Source strength estimation usually accompanies source localization. For anomaly detection, the key is to accurately estimate the background radiation. Then anomaly detection can be performed based on the estimated background radiation. Isotope identification needs radiation spectral information. For source localization and estimation, the key is to build an accurate physical model to represent the radiation transport processes. This section gives a brief review of the existing methods for background radiation estimation, anomaly detection, isotope identification, and source localization.

1.3.1 Anomaly Detection and Isotope Identification

The event that people are interested in, namely if radioactive sources actually exist somewhere, has very low probability to happen. Accordingly, most of the collected data contain only background radiation information. After collecting the data, the first step is to process the data such that only the data consisting of radioactive source information are kept. The highly imbalanced data, in which most measurements are from normal background radiation and only a small portion contains information from radioactive sources, distinguish the radiation detection problem from traditional binary classification problems [15]. Instead, radiation detection is essentially an anomaly detection or outlier analysis problem [16]. Hawkins [17] defined the outlier as "*an observation which deviates so much from other observations as to arouse suspicious that it was generated by a different mechanism*". In nuclear radiation detection field, the data that contain signals from various radioactive sources are naturally defined as abnormal measurements since they originate from processes that are different from those that lead to normal background radiation. The goal of anomaly detection in radiation data is to identify all data that contain information from radioactive sources with a low probability of false positives.

The key to identifying outliers is to distinguish the measurements consisting of information from radioac-

tive sources from those with only background radiation. As will be discussed later, background radiation exists everywhere and it changes with time. The real-world radiation data collected by detectors are noisy. When radioactive source intensity is low or the distance to radioactive sources is large or the radioactive source is well-shielded, the SNR is low, which makes the signals from the actual sources hard to distinguish from the background radiation signals. All these factors make it hard to accurately distinguish between data from radioactive sources and background radiation.

Depending on the data quality, different methods have been developed to identify data from radioactive sources. When detector size is small or data collection time is short (e.g., 1 second), the spectral quality is low. In this case, gross count rate is the only information that can be used to identify anomalies. In the most widely used method, the k-sigma method, some predefined threshold is used to determine whether or not there are radioactive sources. If the count rate of the given data is above the threshold, then it is assumed that some sources have been detected [6, 11]. Although very simple, the k-sigma method works well when the threshold is properly defined and background radiation is correctly measured or estimated. In addition to the k-sigma method, Bayesian estimation and the sequential probability ratio test (SPRT) have also been applied to detect the anomalies. Bayesian algorithms compute the posterior probability distribution based on estimated prior distribution [18]. The performance of these algorithms has been tested using simulations [1, 3, 5, 19]. SPRT is a widely used hypothesis-testing method. It compares current measurements to a previous baseline [11]. SPRT is useful for long-term radiation monitoring [20] and for low-level point source detection and localization with stationary detectors [21, 22].

For cases with only gross count rate data, background radiation is either assumed to be known or can be estimated using the data from surrounding detectors or from historical data. For more modern detectors, however, the collected data may also contain the spectral information, which make it possible to estimate the background radiation purely from single measurement. Several methods have been developed to estimate and separate the background radiation. Kirkpatrick et al. used Poisson rather than Gaussian statistics to compute the background contributions to gamma-ray spectra, which proved to be a sensitive method in the case of small signals in low-count spectra [23]. By making the assumption that the background signal varied smoothly while the signal from radioactive sources varied more rapidly, Fischer et al. proposed the Bayesian method to estimate the background radiation spectra [24]. Morháč et al. proposed a statistics-sensitive non-linear iterative peak-clipping (SNIP) method to determine the peak regions and then eliminated background radiation in the gamma-ray spectra [25, 26]. An efficient iterative filtering method was proposed by Zhu et al. to estimate the background radiation spectra in noisy spectra [27]. Based on the assumption that the local minimum belonged to the background radiation, Alamaniotis et al. proposed a kernel-based machine

learning method to estimate the background radiation for low-count gamma-ray spectra with unknown sources without any prior information [28]. The estimated background spectra from this method had the right shape, but the amplitude of the estimated background spectra was either too high or too low. An algorithm that coupled a robust Kalman filter and an Expectation-Maximization framework was designed by Fraschini and Chaillan to recursively estimate the background spectra [29]. A series of supervised classifiers were also trained through different algorithms, such as support vector machine (SVM) and random forests [11]. The trained classifiers were used to classify new spectra into threat or non-threat categories. These methods need a large quantity of pre-labelled data to train the classifier. They work well for the cases in which only particular sources are included and the testing data are similar to the training data. However, when there are unknown sources in the testing data or the background radiation is changing, correctly labelling the training data is almost impossible, which makes it hard to train a model that generalizes well for unseen spectra. In addition, other methods, such as spectral peak erosion [30], digital filters, and Monte Carlo simulations [31] have also been used in the literature to estimate background radiation.

Spectra are not only useful for anomaly detection, but also important for isotope identification. Spectrum fitting is traditionally used to detect the existence of particular radioactive sources. For the spectrum fitting method, the measured spectrum is fitted with various templates from the pre-calculated library [11]. Spectrum fitting methods can work well when spectra are well-collected, but when data collection time is short (e.g., 1 or 2 seconds) and detector size is small, the collected spectra have low quality and the performance of spectrum fitting methods will be affected. In addition, principal component analysis (PCA) [32, 33] was implemented to extract the information about certain isotopes. In current research, PCA was performed on well-collected data. Its performance on low quality spectra has not been tested. Recently, artificial neural networks have also been applied for isotope identification, especially fully-connected neural networks and convolutional neural networks [34, 35].

Among the proposed methods, a few need to manually establish prior information, some methods require long computation time, and other methods focus on high-resolution spectra [28]. For mobile sensor networks, the detectors usually have small size and the data will be continuously collected only for one to two seconds for each location. This means that the spectra have low-count and low-resolution. Dynamic methods are needed to estimate the background radiation for low-count and low-resolution spectra of unknown sources.

1.3.2 Source Localization and Estimation

Different from anomaly detection, the goals of source localization and estimation are, respectively, to accurately locate the radioactive sources and possibly estimate the strength of radioactive sources. Source

localization and estimation are usually performed based on measured data suggesting the presence of radioactive sources in some parts of the region under surveillance.

To pinpoint the radioactive sources and estimate their strength, gross count rate information is generally used. Chin et al. [21] proposed a two-step approach for the low-level point source identification problem. This approach combined an iterative pruning (ITP) algorithm [36] and a sequential probability ratio test (SPRT) for the N-sensor case. Morelande et al. [37] analyzed the estimation problem for multiple radioactive sources using maximum likelihood estimation (MLE). Liu et al. proposed a background radiation estimation algorithm based on MLE to model the spatial and temporal distribution of background radiation and identify radioactive sources [38, 39]. Zhao and Sullivan analyzed the performance of MLE with grid search on source localization using simulations and experiments [40]. Upasana and Baum extended the work into the three-dimensional case using a two-stage adaptive algorithm based on MLE [41]. Different Bayesian methods have been studied under various circumstances [9, 42–44]. Ram and Verravalli [45] considered the network of spatially distributed sensors and proposed a recursive algorithm to locate and track the diffusing point sources. Hochaum et al. [46, 47] proposed an efficient network flow algorithm to detect the nuclear threats using mobile distributed sensor networks which enabled real-time usage. Simulations were carried out to prove the proposed algorithm's effectiveness. Tandon [8] and Tandon et al. [9] proposed a Bayesian aggregation method to locate the radioactive sources through mobile sensor networks. Their method was based on spectral analysis. Its effectiveness was demonstrated using experimental and simulation results. In addition, other methods, such as mean of estimators (MoE) and mean of measurements (MoM) [48], Markov Chain Monte Carlo (MCMC) sampling [49], and least squares estimation [50] were also implemented for source localization problems.

The above methods are mainly focused on passive search, where radiation detectors are either stationary or randomly moving. Another type of source localization is through active search, in which case radiation detectors will be guided towards the radioactive sources. Ristic et al. [51–53] proposed an information gain driven search technique which combined particle filter and Fisher information gain for locating a radioactive point source. Lin and Tzeng adopted artificial potential field with particle filters to navigate mobile radiation detectors to find radioactive sources in an unknown environment [54]. Ardiny et al. [55] proposed an autonomous exploration strategy based on multi-criteria decision making for radioactive hotspot localization, and the proposed technique was tested using simulation. Recently, a search technique based on reinforcement learning was proposed for radioactive source detection, and the simulation study demonstrated the improvement over the gradient search algorithm [56].

1.4 Overview

There are eight chapters in this thesis, and the structure of the thesis is shown in Fig. 1.2. Chapter 1 outlines the radiation detection problems and describes mobile sensor networks and related work for anomaly detection and source localization. Chapter 2 describes the technical details behind radiation detection and measurements, including nuclear radiation, radiation signal measurement, metrics used to evaluate the radiation detection systems, neural networks, and generally used estimation and optimization techniques. Chapter 3 describes the mobile sensor network platform and software used in this work, as well as the simulation and experiment conducted. Based on the simulations described in Chapter 3, Chapter 4 focuses on the systematic assessment of mobile sensor networks for radiation detection. The influence of several key characteristics are studied. In Chapter 5, Poisson kriging is applied to nuclear radiation estimation and anomalous source detection problems. In Chapter 6, a source localization algorithm based on maximum likelihood estimation is proposed. Simulation and experimental results show that the proposed method can reduce source localization error to be within 3 meters. Chapter 7 presents the exploratory analysis on spectrum-based anomaly detection. A series of machine learning techniques are studied for spectrum-based anomaly detection. Finally, Chapter 8 summarizes the work that has been done in this thesis and gives an overview of future work.

Overall, this thesis systematically evaluates the mobile sensor networks for radiation detection, including system assessment, background radiation estimation, anomaly detection, and source localization. There are several innovations in this work. Through incorporating the geospatial correlation, this work implements Poisson kriging for nuclear radiation estimation, and then extends it for anomalous source detection problems. Further, based on maximum likelihood estimation, this work proposes approaches based on kernel density estimation and grid search to iteratively estimate radioactive source location and strength. Also, this work not only conducts simulations, but also collects experimental data to test the performance of the proposed approaches. Through these near-real data, this work provides more reliable quantifications and analyses. However, there are also some limitations. This work does not consider the radiation spatial variation due to small detector size and short data collection time. There are certain simplifications for simulation and experiments, and the complex shielding effect has not been explicitly considered and studied. Also, this work assumes the radioactive sources are stationary and the scenarios with mobile radioactive sources have not been explicitly studied.

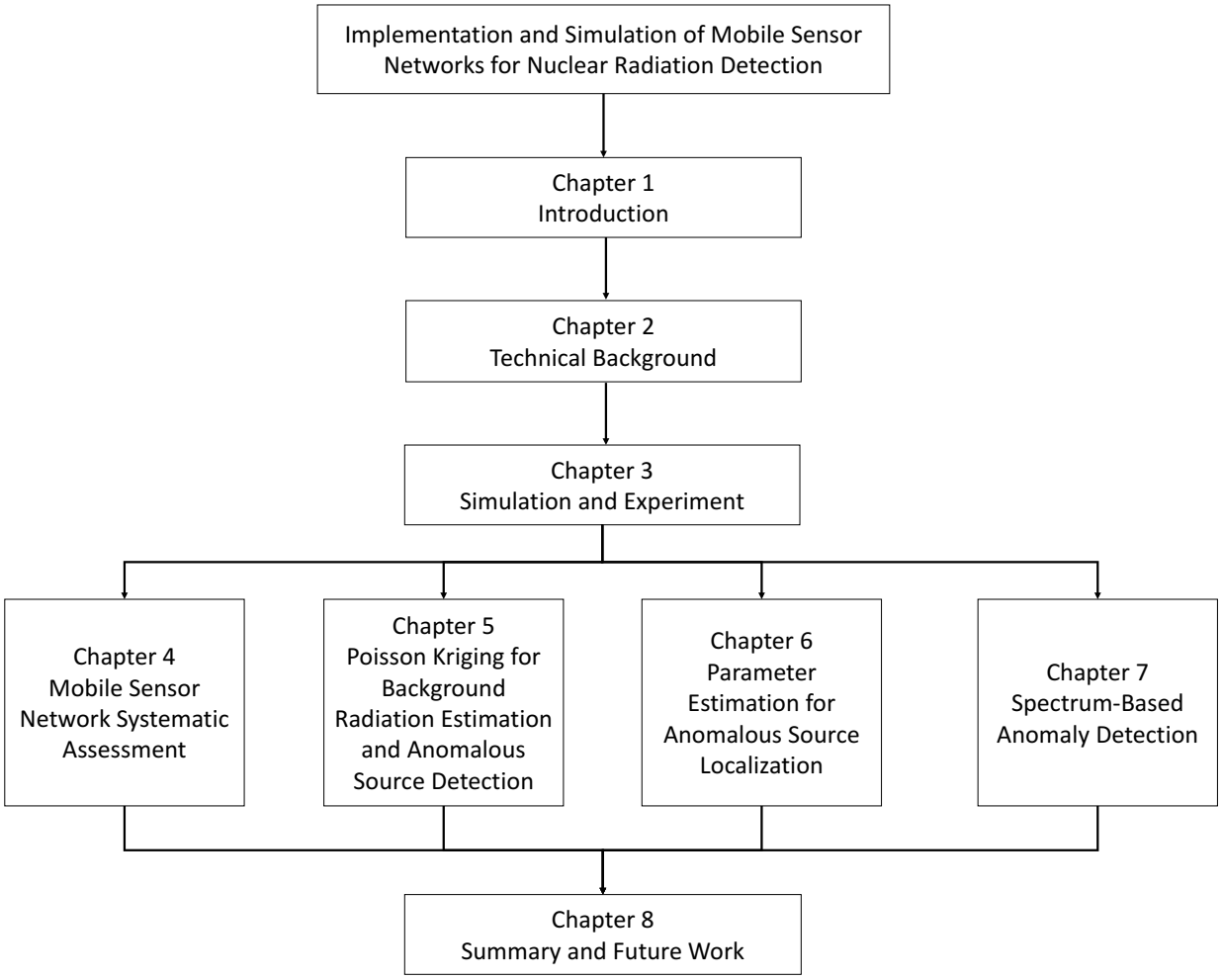


Figure 1.2: Block diagram flowchart of thesis outline.

Chapter 2

Technical Background

2.1 Introduction

A typical nuclear radiation detection system consists of many different aspects. Nuclear radiation, especially radiation from the naturally occurring radioactive materials (NORM), exists almost everywhere. Spectrum and count rate are generally used to quantify the signal strength and radiation characteristics, in which different kinds of radiation detectors play important roles. This chapter briefly describes the technical background behind radiation detection, including background radiation, radiation measurements, metrics used to evaluate the performance of radiation detection system, and some general introduction to mathematical methods used in this thesis such as kernel density estimation, maximum likelihood estimation, optimization techniques, and neural networks.

2.2 Background Radiation

Background radiation generally refers to naturally existing radiation. All humans are continuously exposed to such radiation. Natural radiation sources [57, 58] include:

1. **Cosmic rays** that come from outer space and from the Sun.
2. **Terrestrial radionuclides** that occur within the Earth's crust, in building materials, air, water, foods, and in the human body itself. They mainly come from ^{40}K and the daughter products of the decay series of ^{235}U, ^{238}U, ^{226}Ra, ^{222}Rn, and ^{232}Th [11, 59].

Background radiation varies spatially and temporally. It is influenced by weather conditions, surrounding environment and so on. Studies have shown that the background radiation level increases during precipitation [60, 61]. Soil conditions were also found to affect the background radiation level [62]. Figure 2.1 shows the distribution of background radiation levels at the engineering campus of the University of Illinois at Urbana-Champaign. The data were collected using multiple D3S detectors[1]. The color corresponds to the

[1] Discreet Dual Detector. See Appendix A for technical details of D3S detectors.

background radiation level, where lighter color represents higher background radiation level. The spatial fluctuation of background radiation is clear and some locations have as much as two to three times higher background radiation level than other locations. The background radiation near buildings is usually higher than other places, such as the lawn and parking lot. Three typical high background regions are marked using dashed rectangles. These are the Nuclear Radiation Laboratory (A), the Alma Mater statue (B), and a church (C)[2].

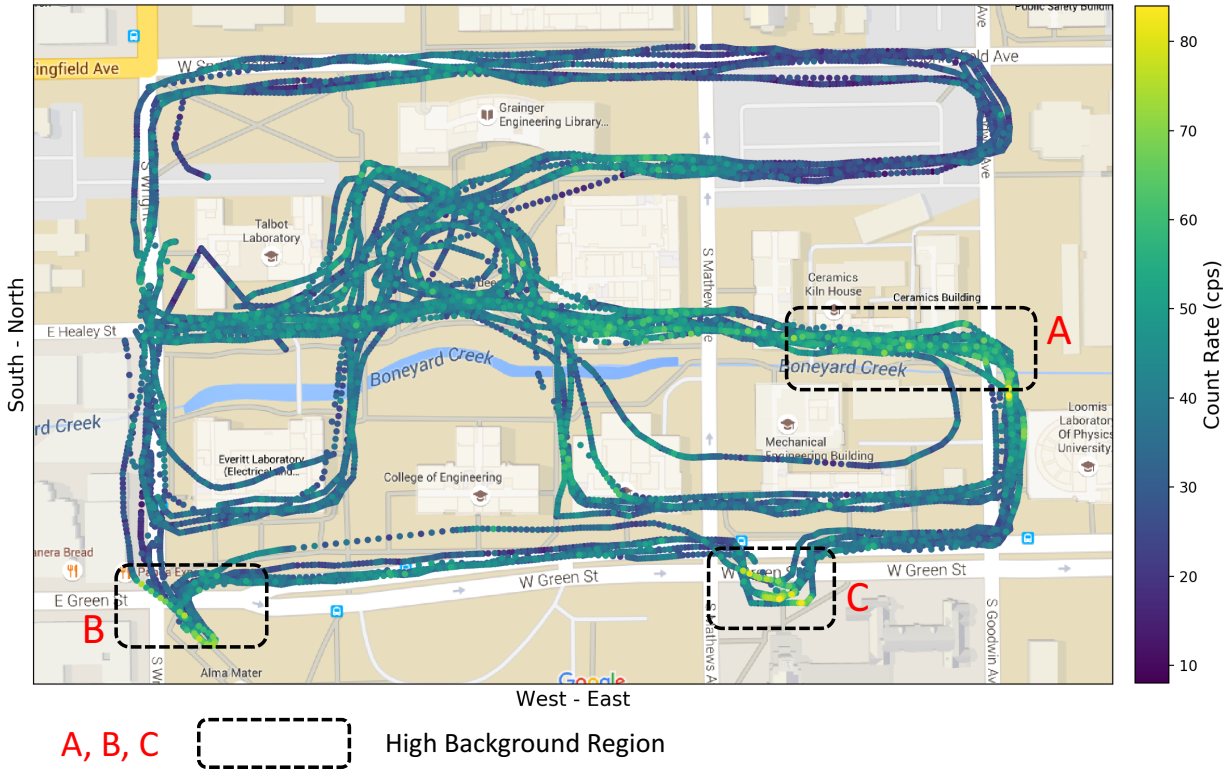
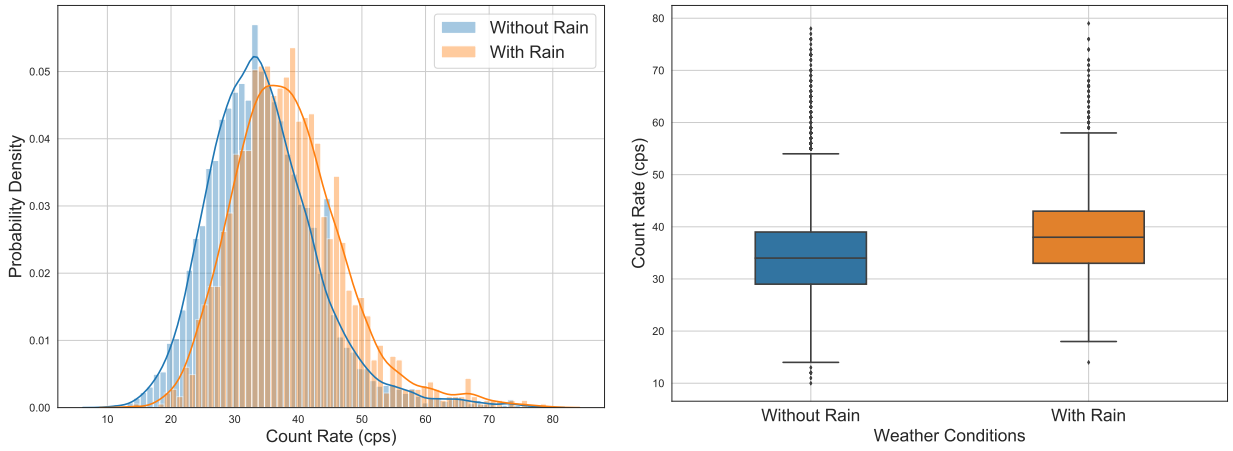


Figure 2.1: Distribution of background radiation on the engineering campus of the University of Illinois at Urbana-Champaign. The data were collected using multiple D3S detectors. Three relatively high background regions, namely the Nuclear Radiation Laboratory (A), the Alma Mater statue (B), and a church (C), are marked using dashed rectangles.

In addition to spatial variation, different weather conditions lead to temporal variation as well. Figure 2.2 shows the influence of rain on the distribution of background radiation. All data were collected in the outdoor environment using D3S detectors. During precipitation, the background radiation level is higher, which is generally assumed due to the ^{222}Rn progeny accretion during precipitation.

[2]For Alma Master status and the church, the high background radiation is due to the materials they contain. For example, the base of the statue and the stairs of the church are both made of granite, which cause the high background radiation.



(a) Histogram of background radiation count rate.

(b) Boxplot of background radiation count rate.

Figure 2.2: Illustration of the change of background radiation under different weather conditions. The data were collected using D3S detectors in outdoor environment.

When the source strength is not strong, the existence of background radiation makes it hard to accurately identify the radiation from radioactive sources, such as the radiation from ^{137}Cs, ^{60}Co, and ^{54}Mn. Accurately estimating the background radiation is the key to addressing many problems in nuclear radiation detection.

2.3 Radiation Measurement

Nuclear radiation needs to be correctly measured to quantify background radiation and identify radioactive sources. There are different types of radiation, such as α particle, β particles, and gamma-ray. This work only considers gamma-ray for nuclear detection. Gamma rays can be measured by various radiation detectors through the interaction of radiation within the detector. Some detectors can only record the gross counts received by the detector, such as a Geiger-Mueller counter. Other detectors can also record the raw spectra, which enables more detailed analyses and application of additional techniques for radiation detection and source identification. Depending on the detectors and the purpose, gross count rate and spectrum are used for different problems.

2.3.1 Gross Count Rate

Gross count measures the radiation counts in some time interval. The raw data collected by radiation detectors are always integers. The count rate is simply the counts in a unit time. In this work, since all radiation data are measured every second, the count rate is also an integer. For radiation detection, the

measured count rate by a detector is generally assumed to follow a Poisson distribution. The probability, p, that a detector collects m counts in a unit time with expectation λ is:

$$p(x = m) = \frac{\lambda^m}{m!} \cdot e^{-\lambda}. \tag{2.1}$$

In Eq. 2.1, λ represents the expected count rate, which stands for the radiation level at certain location. λ consists of the contribution from surrounding background and radioactive sources, in the form of $\lambda \equiv b + s$, where b refers to the radiation from background and s refers to the radiation from radioactive sources. s would be zero when there are no radioactive sources. b is influenced by many factors, such as the surrounding buildings, weather condition and so on. Typically, b is a function of location $\boldsymbol{r}$ and time t. Over a short time interval, b can be assumed to be constant for a given position. s, on the other hand, is mainly influenced by radioactive sources, distance from radioactive sources, and shielding materials between radioactive source and radiation detectors.

Suppose that some radioactive source with source intensity[3] s_0 is placed at location $\boldsymbol{r}_0$, where $\boldsymbol{r}_0 = (x_0, y_0) \in \mathbb{R}^2$, and the i-th detector is placed at position $\boldsymbol{r}_i = (x_i, y_i)$. The distance[4] of the i-th detector to the source is $d_i = ||\boldsymbol{r}_i - \boldsymbol{r}_0|| = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$. Then the average count rate λ_i measured at distance d_i is:

$$\lambda_i = b_i + s_i = b_i + \frac{s_0}{d_i^2} \cdot \exp\left(-\int_{\boldsymbol{r}} \mu d\boldsymbol{r}\right), \tag{2.2}$$

where b_i is the average background radiation count rate at position $\boldsymbol{r}_i$, μ is the linear attenuation coefficient (m^{-1}), and $\exp\left(-\int_{\boldsymbol{r}} \mu d\boldsymbol{r}\right)$ results from the attenuation effect. Table 2.1 lists several common materials and their density as well as the linear attenuation coefficient at different energy levels.

[3]To simplify the mathematical expressions, in this work, source intensity s_0 is defined to be the average count rate measured by a given type of detector that is placed 1 m away from the source without any shielding.

[4]Here, for the distance calculation, only x and y directions are considered. The height of the detector is ignored since it is always assumed to be 1 meter high above the ground.

Table 2.1: Linear Attenuation Coefficient at Gamma-ray Energies of 100, 200, and 500 keV

Absorber	**Density** (g/cm^3)	**Linear Attenuation Coefficient** (cm^{-1})		
		100 keV	**200 keV**	**500 keV**
Air	0.001225	0.000195	0.000159	0.000112
Water	1.00	0.167	0.136	0.097
Carbon	2.00	0.335	0.274	0.196
Aluminium	2.70	0.435	0.324	0.227
Iron	7.87	2.720	1.090	0.655
Copper	8.93	3.800	1.309	0.730
Lead	11.34	59.700	10.150	1.640

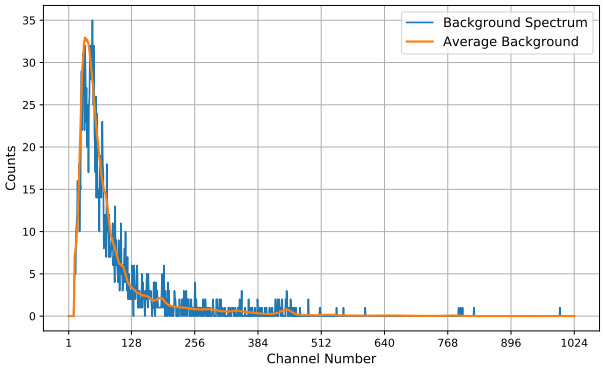
Source: `en.wikibooks.org/wiki/Basic_Physics_of_Nuclear_Medicine/Attenuation_of_Gamma-Rays`.

2.3.2 Radiation Spectrum

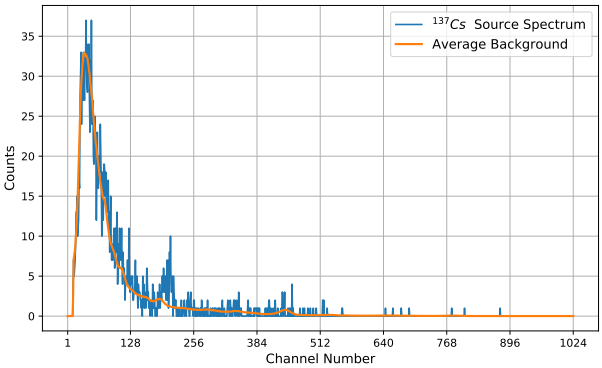
Gross count rate has been widely used for nuclear radiation detection. As will be shown later, it is still a very important quantity for anomaly detection and source localization problems. However, as a scalar value, gross count rate cannot be used to distinguish different radioactive sources. For problems such as isotope identification, a gamma-ray spectrum is needed.

A spectrum is the histogram of gamma-ray counts binned by energies. Different detectors have different energy ranges, but typically, the spectrum contains the records in the energy range of 30 keV to 3 MeV. The number of channels varies according to the detectors. For example, the D3S detector has 512 channels while the Ortec 905 NaI detector has 1,024 channels[5]. A typical background radiation spectrum and the spectra for ^{137}Cs, ^{60}Co, and ^{152}Eu are shown in Fig. 2.3. All spectra are measured using the Ortec 905 NaI detector in room 209D of Talbot Laboratory by the author for 10 seconds. The orange color corresponds to the average background radiation spectrum. Figure 2.3a shows spectra from background radiation. Figure 2.3b, Fig. 2.3c, and Fig. 2.3d show spectra from ^{137}Cs, ^{60}Co, and ^{152}Eu, respectively. The typical gamma-ray energies are listed in Table 2.2 for some frequently used radioactive isotopes.

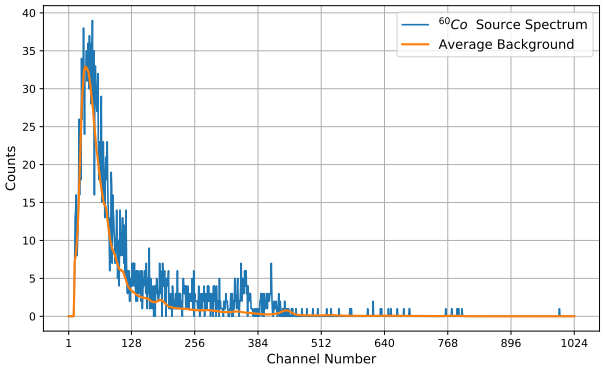
[5]See Appendix A for technical details of the D3S detector and the Ortec 905 NaI detector.



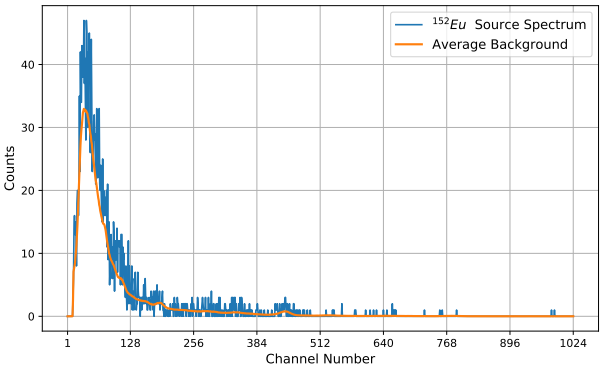
(a) Typical background radiation spectrum.



(b) Typical ^{137}Cs spectrum with background.



(c) Typical ^{60}Co spectrum with background.



(d) Typical ^{152}Eu spectrum with background.

Figure 2.3: Radiation spectra for background radiation, and ^{137}Cs, ^{60}Co, and ^{152}Eu.

Table 2.2: Common Radioactive Isotopes and Major Gamma-ray Energies

Radioactive Isotope	**Typical Gamma-ray Energy (keV)**
^{241}Am	59.5
^{54}Ba	81.0, 276.4, 302.9, 356.0, 383.8
^{57}Co	122.1, 136.5
^{60}Co	1173.2, 1332.5
^{137}Cs	661.6
^{55}Fe	5.9, 6.5
^{54}Mn	834.8
^{22}Na	511.0, 1274.5

2.4 Performance Measurement

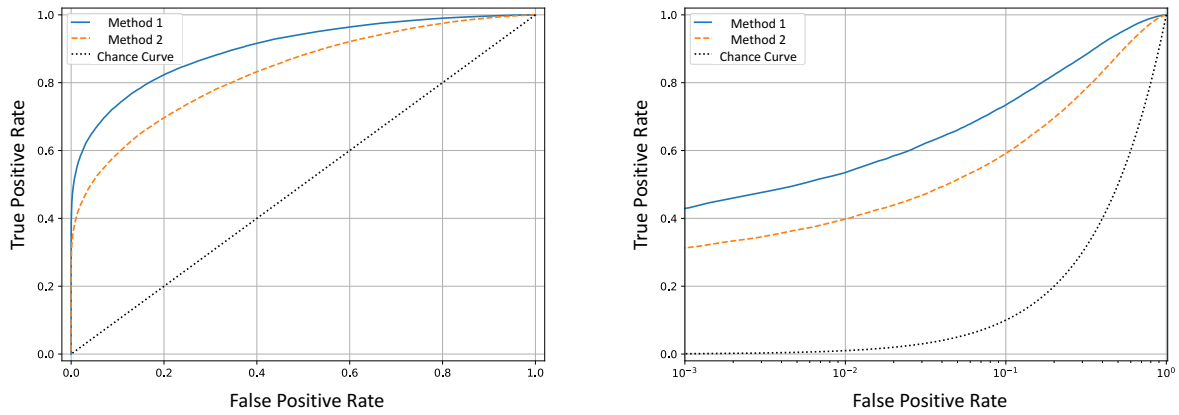
As discussed in Chapter 1, the focus of this work is on anomaly detection and source localization problems. For source localization, the Euclidean distance of the estimated source location from the actual source location is generally used to quantify and compare the performance of source localization algorithms. However, for anomaly detection, traditional metrics used for binary classification problems, such as classification accuracy or error rate, are not appropriate since the proportion of abnormal measurements, indicating the presence of a source, is usually very low. In this work, a receiver operating characteristic curve (ROC curve) and the area under the curve (AUC) [63] are chosen to compare the performance of different anomaly detection techniques.

The ROC curve is the plot of true positive rate versus false positive rate at different thresholds [64]. The definitions of true positive events and false positive events vary among problems. For radiation detection, the true condition has two cases: either there is a radioactive source, or there is no radioactive source. The predictions from the radiation detection system have two results: alert or no-alert. So, there are four outcomes based on the true situation and the predicted results: alert when sources are present, no-alert when sources are present, alert when sources are not present, and no-alert when sources are not present. The corresponding events are defined as: true positive events (TP), false negative events (FN), false positive events (FP), and true negative events (TN), respectively. Based on these definitions, true positive rate (TPR) and false positive rate (FPR) are defined as follows [58, 64]:

$$TPR = \frac{\sum TP}{\sum TP + \sum FN} \tag{2.3}$$

$$FPR = \frac{\sum FP}{\sum FP + \sum TN}. \tag{2.4}$$

A ROC curve shows the change in TPR with increasing FPR. For an ideal system, the true positive rate should be 1 while the false positive rate is 0. The ROC curves for two sample methods are shown in Fig. 2.4a.



(a) Sample ROC curves.

(b) Sample ROC curves in natural log scale.

Figure 2.4: Sample ROC curves. Method 1 and method 2 represent two sample methods in general.

As shown in Fig. 2.4a, ROC curves for two sample methods are plotted along with the chance curve, which represents the performance of random guess. The area under the curve (AUC) indicates the overall performance of different methods. AUC is calculated as the area under the ROC curve with the maximum value of 1. Higher AUC is usually preferred. For example, the AUC of chance curve is 0.5, and the AUC of method 1 is higher than AUC of method 2, in which case method 1 is preferred over method 2.

For radiation detection, since the incidents have very low probability, any radiation detection system should have a low false positive rate to avoid massive wasting of resources. Thus, in addition to overall performance, the likelihood of the system at a given false positive rate is also of concern. But the original ROC curves shown in Fig. 2.4a cannot clearly show the important information when the false positive rate is low. In this case, the ROC curves plotted with the false positive rate in natural log scale is a better choice. The results are shown in Fig. 2.4b, which clearly shows the difference between method 1 and method 2 when the false positive rate is low.

In addition to the ROC curve and the AUC, there are other metrics that can be used, such as precision, recall, and F1 score. Precision is defined as the fraction of true positive events (TP) over true positive events (TP) plus false positive events (FP). Recall has the same definition as true positive rate. Similar to the trade-off between true positive rate and false positive rate, there exists the trade-off between precision and recall, which is usually plotted as the precision-recall curve[6].

[6]See Ref [16] or `https://en.wikipedia.org/wiki/Precision_and_recall` for more details.

2.5 Kernel Density Estimation

Kernel density estimation (KDE) [15, 65] is a nonparametric approach to approximate the probability density function given observed data. For KDE, there is no need to make any assumption about data's distribution. The only need is to choose a smooth non-negative kernel function, and then KDE can give an estimation of the original distribution from which observed data are drawn.

Considering a random 1D sample dataset with N measurements, $x_1, x_2, \cdots, x_N$. Assuming that these samples are drawn from a probability density function $f(x)$, the goal of KDE is to obtain an estimated probability density function $\hat{f}(x)$. Given the kernel function $K(x, x_i)$, the estimated probability density function $\hat{f}(x)$ is:

$$\hat{f}(x) = \frac{1}{C} \sum_{i=1}^{N} K(x, x_i) \tag{2.5}$$

where C is the normalizing constant to make $\hat{f}(x)$ a valid probability density function.

To better illustrate the mechanism behind KDE, a sample case is shown in Fig. 2.5. There are 20 randomly sampled data $x_1, x_2, \cdots, x_{20}$ from a sample Gaussian distribution with mean 0 and standard deviation 1, which are denoted by short blue vertical lines on the figure. For KDE, a Gaussian kernel, $K(x, x_i) = exp(-\frac{(x-x_i)^2}{2h^2})$ where h is the bandwidth, is selected. For each x_i, there is a fitted kernel plotted over the x-axis, which is denoted by the red curves. The estimated probability density, which is the total sum of all kernels, is shown using the green curve. Through KDE, a smooth estimation of probability density is obtained, which can help better understand the distribution of original data.

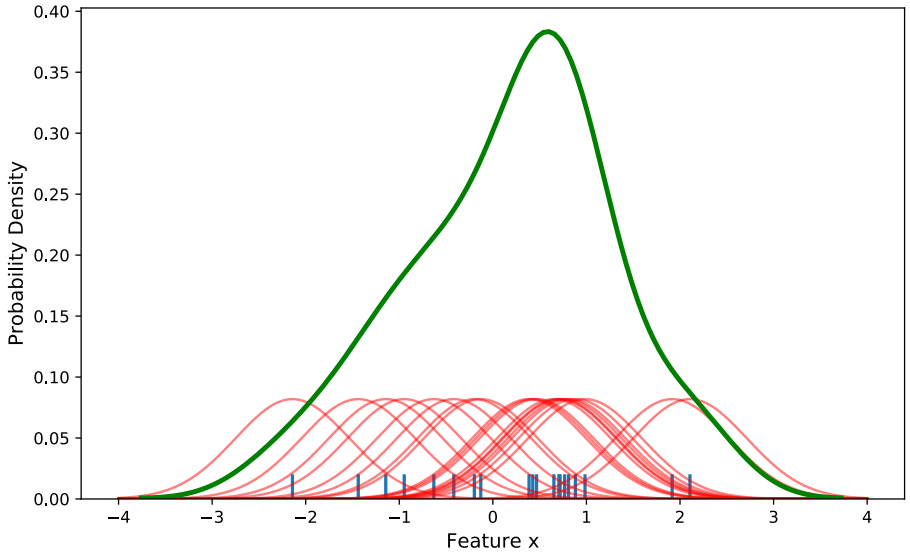


Figure 2.5: Illustration of kernel density estimation.

2.6 Maximum Likelihood Estimation

Maximum likelihood estimation (MLE) is a method used for parameter estimation given observed data. It has been widely used in history [66]. The idea behind MLE is that after formulating the likelihood function using the observed data, the parameters that maximize the likelihood function can be obtained, which are the solution to MLE.

Mathematically, assume that the observed data are denoted by x (e.g., $x_1, x_2, \cdots, x_N$). These data are assumed to be drawn from some probability distribution $f(x;\theta)$ with unknown parameter θ. Then, the likelihood function $\mathcal{L}(\theta;x)$ can be written as:

$$\mathcal{L}(\theta;x) = \prod_{i=1}^{N} f(x_i;\theta). \tag{2.6}$$

Usually, likelihood function $\mathcal{L}(\theta;x)$ is hard to maximize directly. Instead, the log-likelihood function $\ell(\theta;x)$ is much easier to maximize through log-transformation:

$$\ell(\theta;x) = \sum_{i=1}^{N} log(f(x_i;\theta)). \tag{2.7}$$

Then, the solution of MLE is:

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \ell(\theta;x). \tag{2.8}$$

2.7 Optimization Techniques

The maximum likelihood estimation solution in Eq. 2.8 is not always easy to obtain. The log-likelihood function $\ell(\theta;x)$ might not be concave and there could exist multiple local maxima, or the analytical solution could be hard to obtain. For the case in which the log-likelihood function $\ell(\theta;x)$ is concave [67] but the analytical solution is hard to obtain, optimization techniques such as gradient ascent [15], Newton-Raphson [68], and Fisher's Scoring [69] can be utilized.

2.7.1 Gradient Ascent

Gradient ascent (or gradient descent) is the first-order iterative optimization algorithm to find the maximum (or minimum) of a function. Recently, with the development of neural networks [70], gradient ascent and gradient descent have become more popular.

Consider the differentiable concave function $f(\theta)$. In order to find θ that maximizes $f(\theta)$, the iterative scheme for the gradient ascent is:

$$\theta^{(k+1)} = \theta^{(k)} + \eta \nabla_{\theta^{(k)}} f(\theta^{(k)}), \tag{2.9}$$

where k is the number of iterations, $\theta^{(k)}$ represents the parameter estimation after k-th iteration, η is the learning rate, and $\nabla_{\theta^{(k)}}$ is the gradient operator.

For the continuous concave function $f(\theta)$, theoretically, with a gradually decreasing learning rate η, the gradient ascent can find the exact solution that maximizes $f(\theta)$. In practical implementation, however, the iterative process is stopped when there is little change between $\theta^{(k+1)}$ and $\theta^{(k)}$.

2.7.2 Newton-Raphson

In addition to gradient ascent, another choice is the Newton-Raphson, which is a second-order iterative optimization algorithm.

In the case that the target function $f(\theta)$ is concave and differentiable, for the solution θ_0 that maximizes $f(\theta)$, the first derivative of $f(\theta)$ should be 0:

$$\nabla_{\theta_0}(f(\theta_0)) = 0, \tag{2.10}$$

where ∇_{θ_0} is the gradient operator.

To find the solution that makes the gradient to be 0, the second derivative is needed. The iterative solution is:

$$\theta^{(k+1)} = \theta^{(k)} - \Delta^{-1}_{\theta^{(k)}}(f(\theta^k))\nabla_{\theta^{(k)}}(f(\theta^k)), \tag{2.11}$$

where k is the number of iterations and $\Delta_{\theta^{(k)}}$ is the second derivative operator.

After sufficient number of iterations, Eq. 2.11 can approximate the solution that maximizes the target function $f(\theta)$.

2.8 Neural Networks

Recently, neural networks [70], especially deep learning models, have been widely used for many problems, such as computer vision and natural language processing. This section gives a brief description of feedforward neural networks.

Neural network is a mathematical model that tries to approximate some function $y = \hat{f}(x)$, where x is the input and y is the output. A sample feedforward neural network is shown in Fig. 2.6.

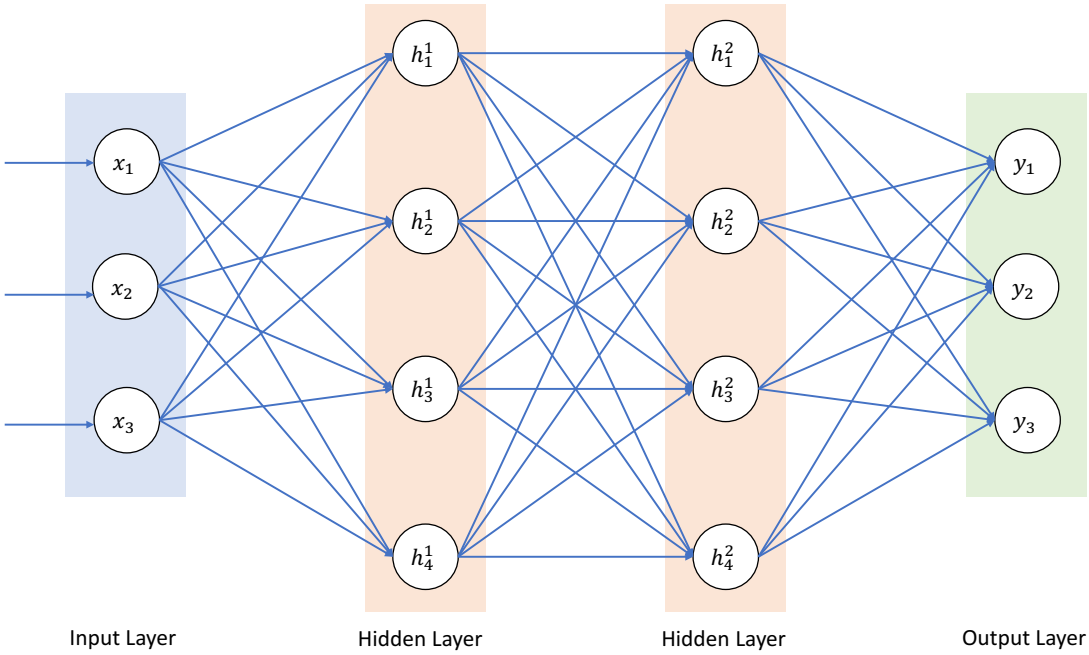


Figure 2.6: Illustration of neural network.

This network has one input layer (x), two hidden layers (h^1 and h^2), and one output layer (y). For this sample neural network, there are three input nodes (x_1, x_2, and x_3). For the first hidden layer, there are four hidden nodes (h_1^1, h_2^1, h_3^1, and h_4^1). For the second hidden layer, there are four hidden nodes (h_1^2, h_2^2, h_3^2, and h_4^2). And for the output layer, there are 3 nodes (y_1, y_2, and y_3). Except the input layer and output layer, the number of hidden layers and number of nodes at each hidden layer are hyperparameters that need to be decided by the user.

For each hidden layer and output layer, the input is the linear combination of previous layer's output. For example, for the j-th node in the k-th layer[7], the input is $\sum_{i=1}^{m} w_{ij}^k x_i^{k-1}$, where there is assumed to be m nodes in layer $k-1$. Then, the output of j-th node in the k-th layer is $a(\sum_{i=1}^{m} w_{ij}^k x_i^{k-1})$, where $a(\cdot)$ is a non-linear activation function. This process can be better illustrated using Fig. 2.7, where the input and output of the first node in the first hidden layer is clearly shown.

[7]Here, the input layer can be seen as the 0-th hidden layer.

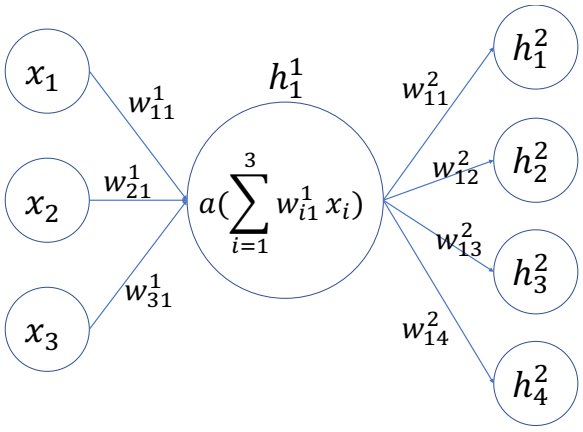


Figure 2.7: Illustration of neural network single node operation.

Through defining loss function, for example, mean squared error for regression problems or cross entropy for classification problems, the neural network can be trained by minimizing the loss using the mechanism called backpropagation [71]. The non-linear activation function is one of the key factors that give neural network the capability to model complex functions. There are several commonly used non-linear activation functions, such as sigmoid function, tanh function, and rectified linear unit (ReLU). The mathematical formats are listed as follows:

$$\text{Sigmoid function:} \qquad a(x) = \sigma(x) = \frac{1}{1+e^{-x}} \tag{2.12}$$

$$\text{Tanh function:} \qquad a(x) = tanh(x) = \frac{(e^{x}-e^{-x})}{(e^{x}+e^{-x})} \tag{2.13}$$

$$\text{ReLU function:} \qquad a(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}. \tag{2.14}$$

The above feedforward neural network is one of the most simple neural network structure. In addition to feedforward neural network, there exist other neural networks, such as convoultional neural networks (CNNs) [72, 73] and recurrent neural networks (RNNs) [71]. In this work, only feedforward neural work is implemented for spectrum-based anomaly detection.

2.9 Summary

Understanding the difference between background radiation and the radiation from SNM is important for understanding the goals of this work, and understanding the signal collection process can help determine the appropriate signals to be used for specific tasks. The performance metrics are important to evaluate and compare different algorithms. Also, choosing the correct optimization techniques and algorithms is the key to solving these problems. This chapter has given a brief overview of the technical background that is important for this work. In the next chapter, the simulation and experiments conducted in this project are introduced, which provide the data used in this project.

Chapter 3

Simulation and Experiment

3.1 Introduction

Previous chapters defined the problem to be solved and gave technical background information needed for this project. To test the techniques for nuclear radiation detection using mobile sensor networks, real-world experiments would be ideal. However, large-scale and long-time deployment of mobile sensor networks in a large urban environment is currently not feasible because of limited available resources. In addition, due to safety concerns, it is hard to conduct experiments with actual radioactive sources. Based on these reasons, simulation plays an important role in the evaluation of techniques used for nuclear radiation detection with mobile sensor networks.

This chapter briefly describes experiments and simulations conducted in this work. More specifically, section 3.2 describes the hardware and software used for experiments and synthetic data generation, section 3.3 discusses the simulations conducted for mobile sensor networks, and section 3.4 describes the details of the experiment and simulated source injection process. The measured background radiation data and simulated radioactive source data are used throughout this work.

3.2 Mobile Sensor Network Platform

3.2.1 Mobile Sensors

Mobile sensors play the most important role in mobile sensor network systems. The sensors should be able to collect GPS location, radiation level, timestamps, and possibly weather conditions in real time. Collected data should be saved and made available for analysis in real time. Data will be even more useful if data from all mobile sensors are available for analysis simultaneously, in real time.

However, most available detectors themselves do not have the capability to transmit the collected information to a central location. Thus, each sensor should contain two parts: a gamma-ray detector to collect radiation data, and a second device to collect and transmit radiation and location information. In this work,

a smart phone is used to transmit such information. A picture of a mobile sensor containing a D3S detector and a smart phone is shown in Fig. 3.1.

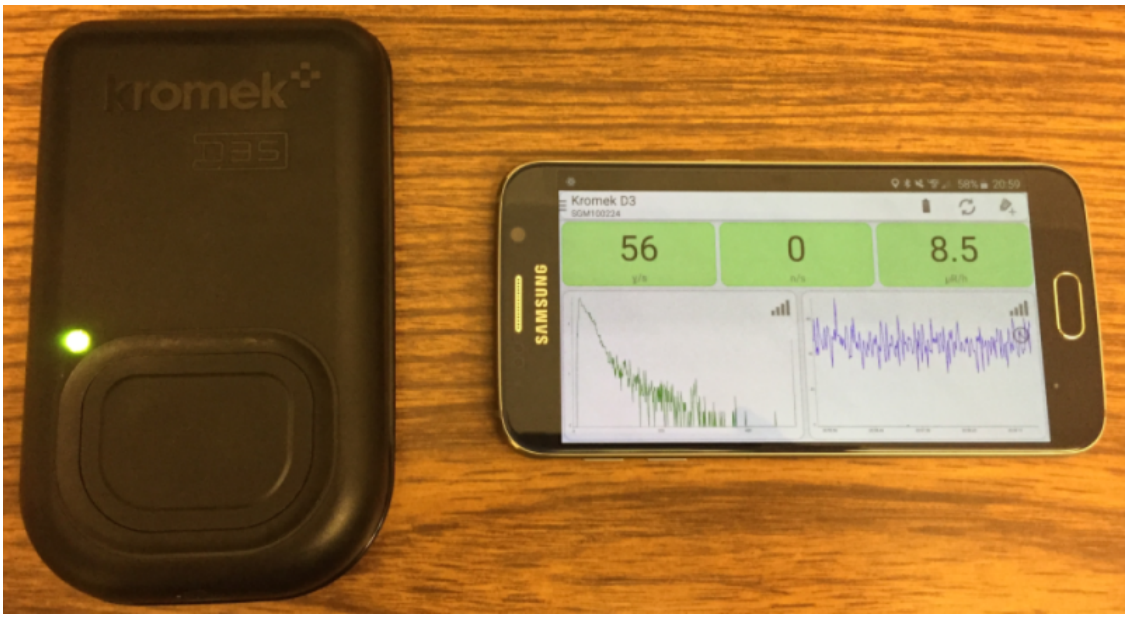


Figure 3.1: D3S detector and cell phone as single mobile sensor.

The small detector used in this work is from Kromek Group plc, which is called the Discreet Dual Detector, or D3S detector. D3S detector is a hybrid gamma/neutron detector used for national security applications. As a 132 $mm \times 80\ mm \times 23.5\ mm$ CsI(Tl) detector for gamma-ray detection, D3S detector's gamma-ray energy range is from 30 keV to 3 MeV. The detector is Bluetooth enabled. It can collect data for every second and then transmit the data wirelessly to the smart phone through Bluetooth. The phone used in this work is a Samsung Galaxy S6 smart phone. It is used to record the corresponding GPS location information and timestamps. A mobile software application was also developed by the research group [4, 38, 74, 75] to connect the smart phone and the Amazon cloud. Through the application, the D3S detectors and smart phones can collect data every second and then transmit the data into cloud with some latency (e.g., around 10 minutes using current application [75]), which enables timely processing and analysis. For more information about the radiation detectors used in this work, see Appendix A.

3.2.2 Synthetic Data Generation

The portable mobile sensors described previously make it easy to collect and transmit data. However, due to safety concerns, actual radioactive sources are not used in this work. Instead, synthetic data are simulated

using GADRAS software[1]. GADRAS can simulate the background radiation for different locations as well as the radiation from various isotopes, such as ^{54}Mn, ^{137}Cs, and ^{60}Co. Its performance has been validated by a series of benchmark sources [78] and analyzed by Stinnett [79]. Two types of detectors, D3S detector and Ortec 905 NaI detector[2], are simulated for this work.

The simulated data contain two parts, radiation from background and radiation from radioactive sources. For background radiation, to better reflect the spatial variation, the background radiation is simulated based on radiation levels in Chicago and Salt Lake City, covering both low background (Chicago) and relatively high background radiation (Salt Lake City) cases. According to the simulated data from GADRAS, the average and standard deviation of background radiation level in Chicago for D3S detectors is 31.3 *cps* and 5.6 *cps*, and they are 38.6 *cps* and 6.2 *cps* for Salt Lake City. For radiation from radioactive sources, the radiation field is simulated within a radius of 20 m around the source to reflect the influence of distance on radiation. A schematic diagram of the simulation process for radioactive sources is shown in Fig. 3.2. In the simulated region, multiple spectra are simulated for the future sampling purpose at discrete distances (for example, every 0.5 m) from the source. For more details on the synthetic data simulation process and how to use the simulated data, please see Appendix C and Appendix D.

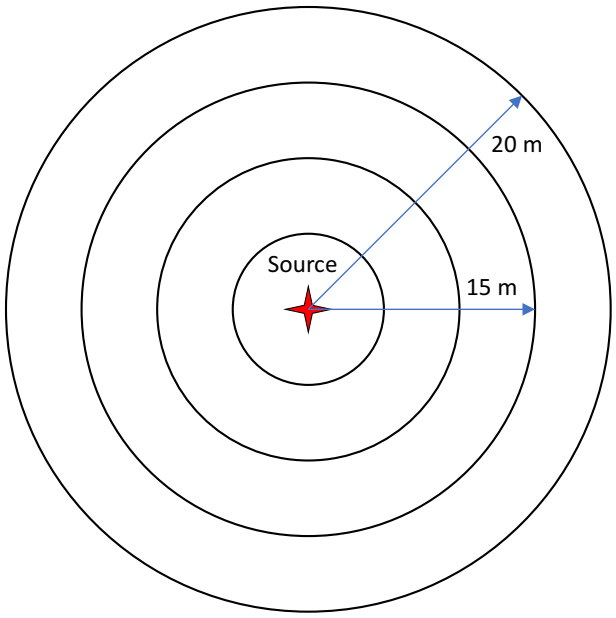
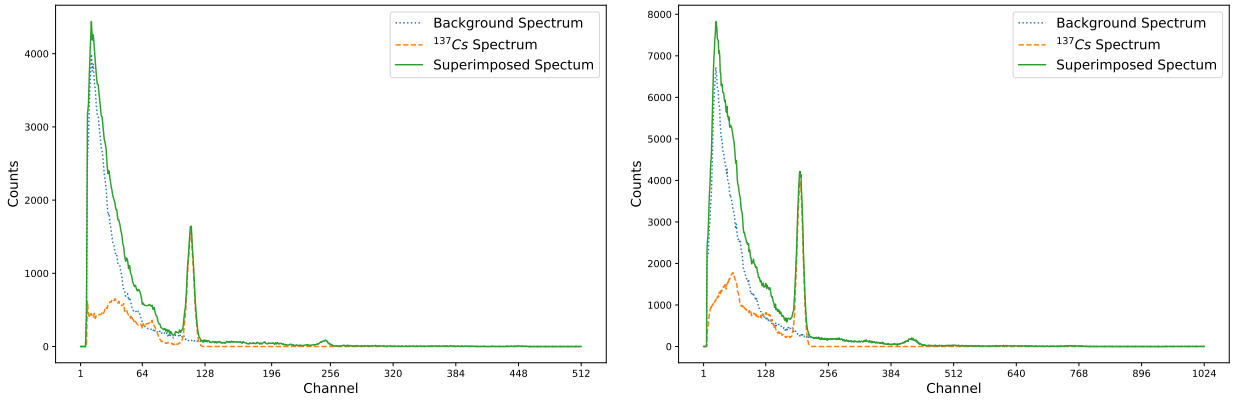


Figure 3.2: Schematic diagram of radioactive source spectrum simulation process. The spectra are simulated at different distances from the radioactive source.

[1]Gamma Detector Response and Analysis Software (GADRAS) is a spectrum simulation and analysis software developed and maintained by Sandia National Laboratories [76, 77]. For more details about GADRAS, see Appendix B.

[2]See Appendix A for technical details of D3S detector and Ortec 905 NaI detector.

Once the background radiation spectrum and the radioactive source spectrum are simulated or collected, based on the assumption that the radiation from background and radioactive sources are two independent processes, the spectrum that will be measured by a detector is simply the sum of two spectra. This process is illustrated in Fig. 3.3, where the spectra for background radiation and ^{137}Cs source with 1 μCi source intensity as well as the combined spectrum are shown for the D3S detector (Fig. 3.3a) and Ortec 905 NaI detector (Fig. 3.3b).



(a) D3S detector.

(b) NaI detector.

Figure 3.3: Illustration of synthetic data superposition process. The background and radioactive source spectra are simulated using GADRAS for D3S detector and NaI detector. The 1 μCi ^{137}Cs source is placed 1 meter away from the detector. All spectra are simulated for 1 hour.

To study the influence of source intensity on the performance of mobile sensor networks, the radiation propagation from ^{137}Cs with increasing intensities (100 μCi, 300 μCi, 500 μCi, 1,000 μCi, and 2,000 μCi) are simulated for the D3S detector and Ortec 905 detector[3]. The simulated source intensities in counts per second (*cps*) for different detectors are shown in Table 3.1.

[3]In this work, since most part only uses the count rate information, the exact isotope category is not very important. The 5 different source intensities are selected to cover the scenarios from very low source intensity (e.g., 100 μCi) to relatively strong source intensity (e.g., 2,000 μCi). These analyses can be easily extended to the cases with stronger shielded radioactive sources.

Table 3.1: Simulated Source (^{137}Cs) Intensity in Counts Per Second for D3S Detector and Ortec 905 Detector

Source Intensity (μCi)	**D3S Detector** (cps)	**Ortec 905 Detector** (cps)
100	127	735
300	380	2,171
500	626	3,564
1,000	1,220	6,842
2,000	2,354	12,620

Note: source intensity is defined to be the average count rate measured by a given type of detector that is placed 1 m away from the source without shielding effect. The unit is counts per second (cps).

3.3 Simulation

In this work, based on the synthetic data simulated using GADRAS, a simulation platform is developed to model the scenario in which multiple radiation detectors randomly move around the city. A simulated city map with multiple mobile sensors is shown in Fig. 3.4. The map is segmented into a 6×6 grid of 36 blocks. Each block represents one city block. The size of each block is $100\ m \times 80\ m$ and the width of roads is 10 m. Mobile sensors are denoted by (blue) rectangles while four fixed radioactive sources are denoted by (red) circles.

In the simulation, 50 detectors are assumed to randomly move only along the roads (only 16 are shown in Fig. 3.4), and they are restricted to move along the roads only. Each detector is assumed to have a constant speed, which is sampled from uniform distribution in the range of 1.4 m/s to 1.8 m/s. When the detector moves to the intersection of roads, it has equal probability to move along the original direction, turn left, or turn right. It is assumed to not make "U" turn at the intersection[4]. When the detector reaches the boundary of the simulated area, it is assumed to turn around and the simulation continues. In other cases, the detectors are assumed to only move along the original direction. Radioactive sources are placed at different locations. Four typical source locations[5] are shown in Fig. 3.4, where source 1 and source 2 are placed in the middle of the roads, while source 3 and source 4 are placed 5 m off the roads. For scenarios corresponding to sources 1 and 2, since detectors can only move along the roads, the shielding effect of the surrounding buildings is

[4]In real cases, "U" turn might happen and the detector may have higher probability to go straight than turning left or turning right. Here, there are some simplifications for detectors' movement. But this kind of simplification has little influence on the conclusion.

[5]These 4 cases are used to represent the typical source locations. In the actual simulation, there are multiple source locations for each case. For example, for scenario 2 where source is placed at the intersection, there are multiple simulations where simulated sources are placed at different intersections.

ignored. But for scenarios corresponding to sources 3 and 4, the shielding effect needs to be considered. In this work, given that the simulated source is relatively weak, the linear attenuation coefficient is assumed to be 0.15 m^{-1} for the surrounding buildings, which is around 10 times the linear attenuation coefficient in the air.

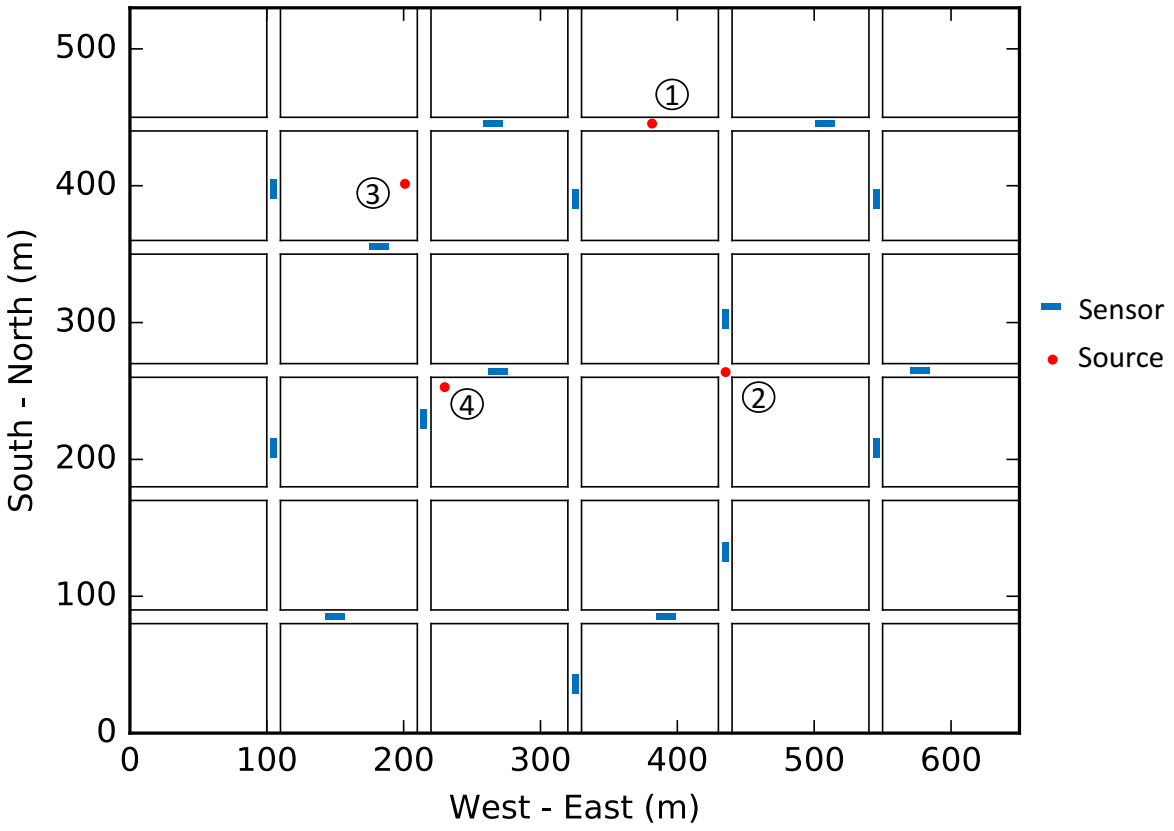


Figure 3.4: Illustration of simulated city map.

For each detector, radiation data and location information are collected every second. The collected data include detector ID, timestamps, exact location information, background radiation count rate, source radiation count rate, and possibly the whole spectrum. The background radiation count rate is sampled from the simulated database described in section 3.2.2 based on its location. For example, when the detector moves to the low background radiation region, its count rate is sampled from the simulated data based on Chicago. When it moves to the high background radiation region, its count rate is sampled from the simulated data based on Salt Lake City. For the count rate from radioactive sources, if the detector is far away from radioactive source (e.g., more than 20 m away), the value would be 0. If the detector is close to the radioactive source, its distance to the source is calculated and the corresponding source count rate is sampled from the simulated radioactive source data at the corresponding distance. This process is repeated

for every second during the simulation[6].

The simulated background radiation map is shown in Fig. 3.5. To better reflect the spatial variation of background radiation, the lower left part of the simulated area is set to have a higher background radiation level. The high background radiation region and low background radiation region can be roughly separated by the dashed curve on the figure. The radius of high background radiation region is around 400 meters. The background radiation data simulated based on Salt Lake City are used to represent the high background radiation, and the simulated radiation data from Chicago are used to represent the low background radiation. Figure 3.5 shows the average background radiation level. Point values are estimated using the measurements from surrounding 30 meters region. The average count rate at position A is 38.5 *cps*, while the average count rate at position B is 31.3 *cps*.

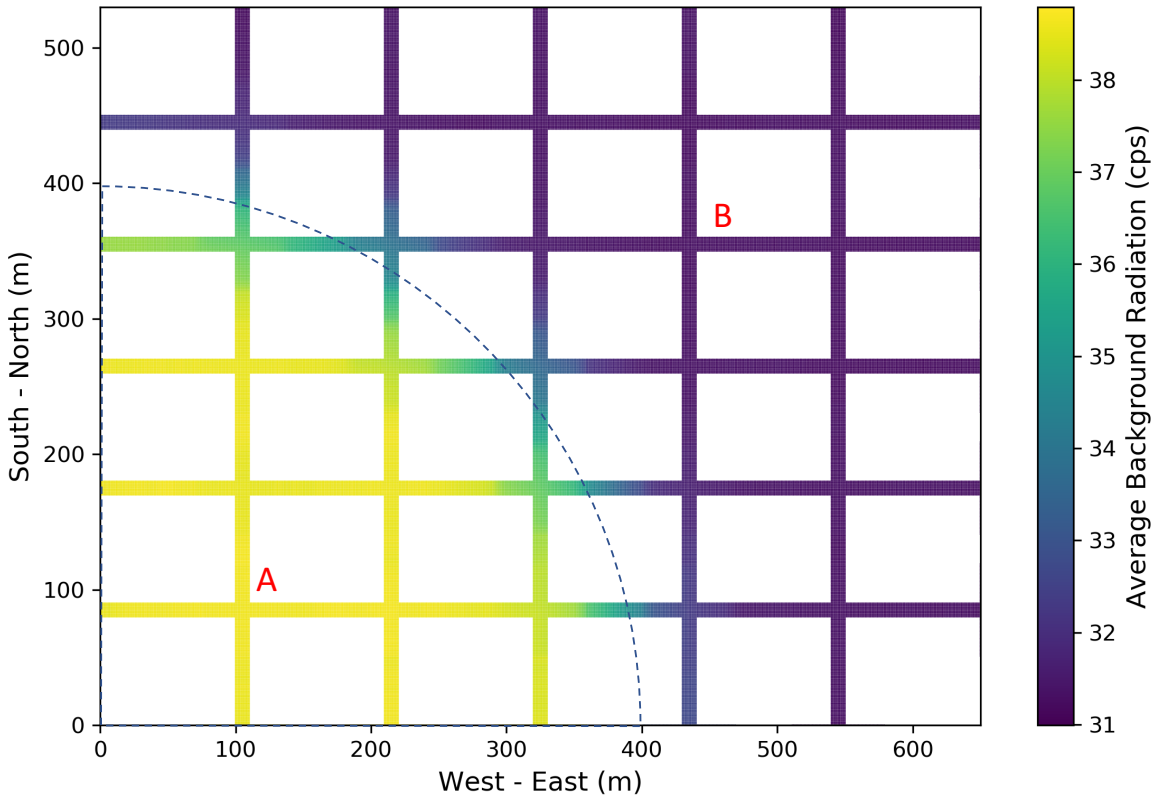


Figure 3.5: Illustration of simulated background radiation. Radiation data are simulated using GADRAS for D3S detectors. Higher background radiation region is located in the lower left part.

[6]See Appendix E for more details about the simulation process.

3.4 Experiment and Source Injection

In addition to simulation, an experiment with a small mobile sensor network system [4, 38, 74, 75] was conducted by the author on the engineering campus of the University of Illinois at Urbana-Champaign to test the proposed algorithms in near-real cases.

The experiment was conducted in an outdoor region of approximately 500×400 m^2 in area. During the experiment, operators who carried D3S detectors walked at the normal speed (around 1.4 m/s) in the experimental area. The experimental area and the walking paths are shown in Fig. 3.6. The operators walked along the designated paths for multiple times. D3S detectors were placed in their pockets. The recorded data include the detector's ID, latitude, longitude, radiation counts per second, and the corresponding timestamps[7]. More detailed description of experimental procedures and collected data are included in Appendix F.

Figure 3.6 shows locations[8] of each measurement and the background radiation level in the experimental area. Clearly, background radiation level increases near the buildings. Three typical high background radiation regions are denoted using dashed rectangles, A, B, and C. The background radiation level at other places, such as the lawn and parking lot, is lower.

[7]The recorded location information might have certain errors. In this work, the recorded location is treated as the actual location and source injection is based on the recorded location information. Thus, the GPS error has little influence on this work.

[8]The original data only have longitude and latitude information. By setting location with longitude equals -88.229402 and latitude equals 40.109666 as the origin, the locations of measurements and buildings can be transformed (using meter) according to their locations relative to the origin.

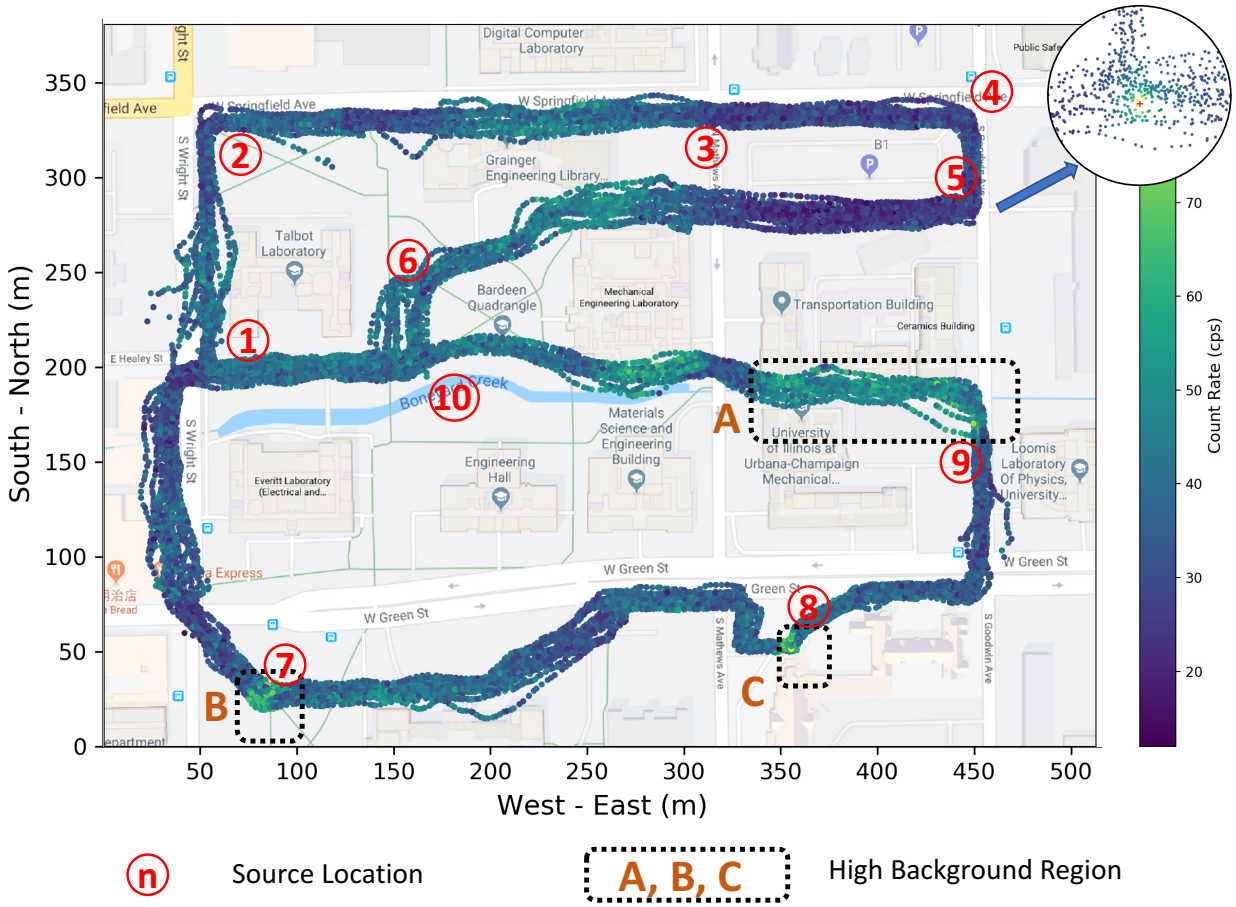


Figure 3.6: Illustration of experimental area, background radiation measurements, radioactive source locations, and source injection process. Three relatively high background regions are denoted using dashed rectangles A, B, and C.

During the experiment, the data were collected in the morning and afternoon by the author, and there were no radioactive sources placed in the experimental region. After collecting the background radiation data, next step is to inject the simulated radioactive source into the measured background data to obtain the "injected" experimental data. Here, the radiation data collected in the morning are used to model the background radiation and the radiation data collected in the afternoon are used to inject simulated radioactive sources. Using GADRAS, the source information of ^{137}Cs with 5 different intensities (127 *cps*, 380 *cps*, 626 *cps*, 1,220 *cps*, and 2,354 *cps*) were simulated. These radioactive sources were injected into the experimental region at 10 different randomly selected locations, which are denoted by numbers inside the circle (from 1 through 10) in Fig. 3.6. To test the performance of the proposed algorithm under different cases, the radioactive sources were injected at 3 different distances (1 m, 5 m, and 10 m) from the walking paths. So, for each injected source location, there are 15 different scenarios in total (5 different source

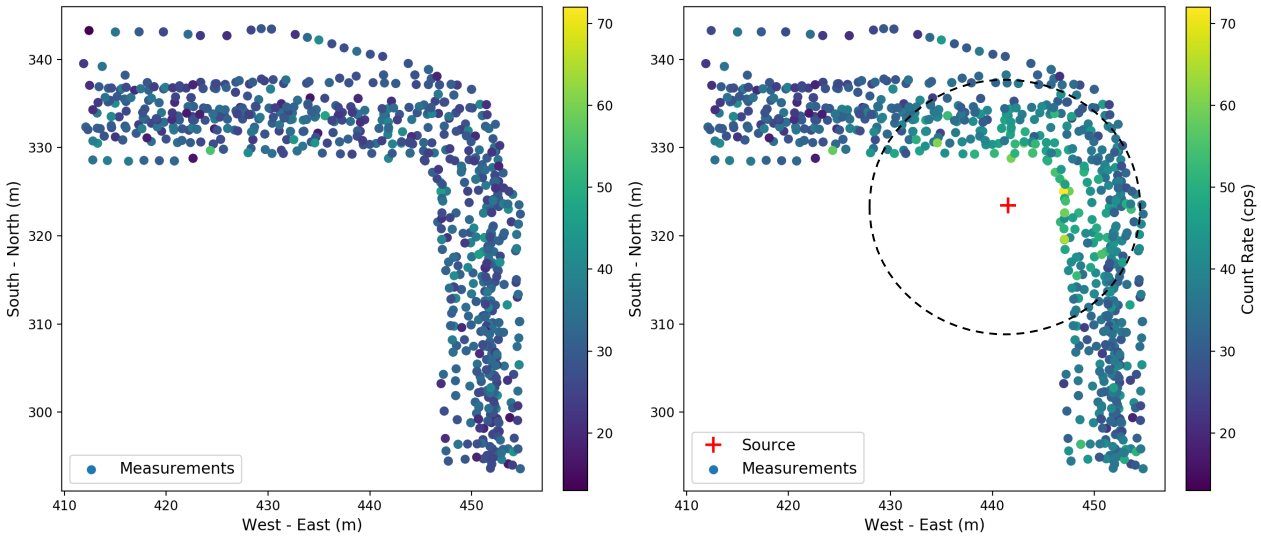
intensities and 3 different distances). The detailed locations of injected radioactive sources are listed in Table 3.2. Each column represents the location of one source injection location, and rows show different injection distances, from 1 m to 10 m.

Table 3.2: Relative Locations of Injected Radioactive Sources

Configuration		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1 m away	**x dir. (m)**	55.9	54.8	310.7	446.8	450.6	172.7	86.5	360.9	453.7	167.9
	y dir. (m)	196.9	328.8	330.6	331.4	281.7	245.4	23.6	50.5	179.1	204.5
5 m away	**x dir. (m)**	59.9	58.6	306.7	443.0	447.2	176.1	86.5	358.1	450.4	171.4
	y dir. (m)	200.5	324.8	326.6	328.4	285.7	243.4	27.7	53.3	176.6	206.2
10 m away	**x dir. (m)**	65.0	63.7	301.8	441.5	442.3	180.4	86.6	354.9	446.5	175.4
	y dir. (m)	205.2	320.0	321.2	323.5	290.8	241.5	32.7	56.8	173.8	209.2

The relative source location is calculated by the distance to the origin, which has longitude equals -88.229402 and latitude equals 40.109666. The location is transformed into meters in both x and y directions.

After injecting the simulated radioactive source, the measured radiation count rates from nearby locations are influenced. This process is shown in Fig. 3.7. When the simulated source is injected, the measured count rates within certain distance from the injected source would increase. For example, the count rates within the dashed circle in Fig. 3.7b are increased by a certain level determined by the distance from the injected sources and the source intensity. For the locations that are far away from the injected sources, there is almost no influence on radiation count rate.



(a) Before source injection.

(b) After source injection.

Figure 3.7: Illustration of synthetic source injection process. After source injection, the radiation level around the injected sources is increased.

3.5 Summary

This chapter briefly describes the hardware and software used for experiments and simulations, detailed experimental procedures conducted using a small mobile sensor network system, and the simulation platform built for this project. In the following chapters, the experimentally collected and simulator generated data are used to evaluate the proposed methods for anomaly detection and source localization.

Chapter 4

Mobile Sensor Network Systematic Assessment

4.1 Introduction

There are many factors that influence the performance of mobile sensor networks, such as radiation detectors, number of detectors, radiation detection algorithms and so on. It is important to understand the influence of these factors on the performance of mobile sensor networks. This chapter first gives an assessment of mobile sensor network system using a modified k-sigma method for anomaly detection. The influence of detector characteristics and the number of detectors is then discussed. The simulated data obtained from the simulation platform as described in Chapter 3 are used to quantify the influence.

4.2 Regional Detection

As mentioned in Section 1.3, various anomaly detection techniques for radiation detection have been proposed and studied. Traditional anomaly detection methods are mainly based on gross count rate. There is increasing interest in spectrum-based anomaly detection techniques.

One of the most widely used methods is the so-called k-sigma method. The traditional k-sigma method uses the deviation of the measurement from the estimated background radiation as the anomaly score [11,38]. Equation 4.1 gives the mathematical definition of anomaly score for the k-sigma method:

$$AS_{k\sigma}(x) = \frac{x - \mu}{\sigma}, \tag{4.1}$$

where $AS_{k\sigma}$ stands for the anomaly score of the k-sigma method, x is the measured gross count rate, μ is the background radiation level, which is usually estimated using the average of surrounding measurements, and σ is the estimated background radiation standard deviation.

The anomaly score calculated using the k-sigma method has no limit on its range, which means that the anomaly score can be very large. It is desirable to define an anomaly score that is bounded within a range, so that the ROC curve and AUC can be better calculated. Suppose the desired range is from 0 to 1,

through normalization by dividing the maximum probability $Poi(\lfloor x \rceil; \mu)$, the calculated Poisson probability $Poi(x; \mu)$ can be rescaled into the range of 0 to 1. This leads to a new Poisson anomaly score definition:

$$AS_p(x) = \begin{cases} \frac{1}{2} \frac{Poi(x;\mu)}{Poi(\lfloor x \rceil;\mu)} & x \leq \mu \\ 1 - \frac{1}{2} \frac{Poi(x;\mu)}{Poi(\lfloor x \rceil;\mu)} & x > \mu \end{cases}, \tag{4.2}$$

where $Poi(x; \mu)$ refers to the probability calculated from Poisson distribution as defined in Eq. 2.1, $\lfloor x \rceil$ is the nearest integer of x, and μ is the estimated background radiation. The anomaly score defined in Eq. 4.2 is bounded over the range of $[0, 1]$. Variation of Poisson anomaly score $AS_p(x)$ as a function of k-sigma anomaly score $AS_{k\sigma}(x)$ is shown in Fig. 4.1.

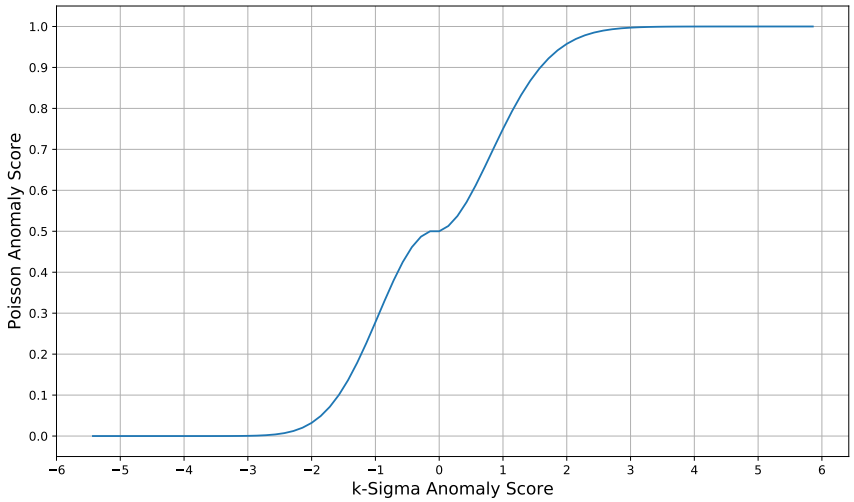
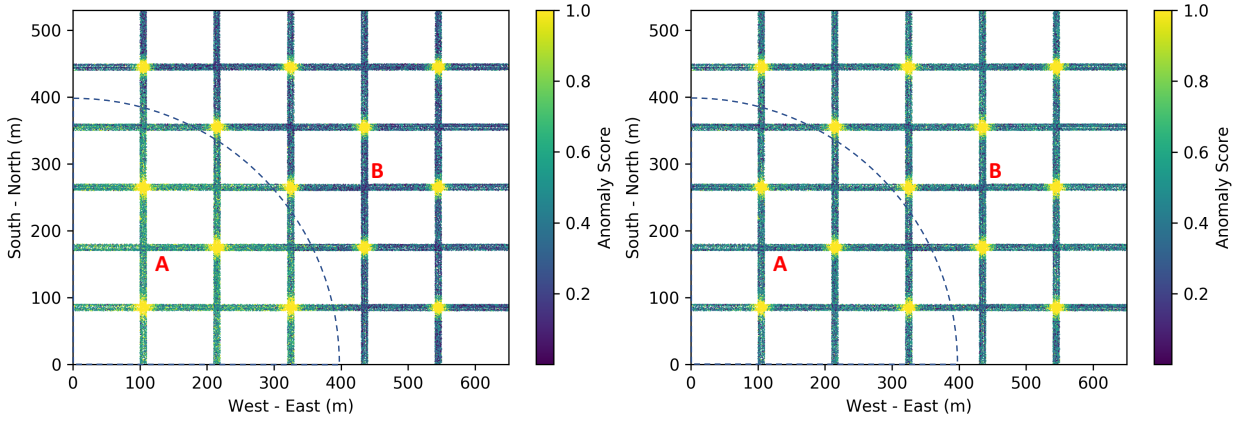


Figure 4.1: Relationship of Poisson anomaly score and k-sigma anomaly score.

The estimation of μ is the key to successfully characterizing abnormal measurements. With stationary detectors, this task can be accomplished using the historical measurements from single detector. With mobile sensor networks, however, it is challenging to get continuous measurements from fixed positions. Two strategies can be applied to address this challenge. The first strategy is to use the measurements across the entire region to estimate the background radiation at the desired location. The anomaly score from this method is named the "global" anomaly score. On the other hand, the background radiation can also be estimated from a small region around the point of interest. The anomaly score from this method is named the "regional" anomaly score or the "local" anomaly score. For the global anomaly score, it is easy and efficient to retrieve all the data to estimate the background radiation. But it ignores the spatial variation of background radiation. On the other hand, the regional anomaly score needs to estimate background radiation level μ for different locations, which is less efficient, especially when the historical data volume is

large (as will be discussed in Chapter 5, certain methods, such as Poisson kriging, can be used to estimate the background radiation level for locations where no measurements were collected).

To compare the performance of global and regional anomaly detection techniques, the simulation platform described in Chapter 3 is used to simulate the scenario where 50 mobile D3S detectors randomly move in the experimental region. Each detector is assumed to have a constant speed, which is sampled from uniform distribution in the range of 1.4 m/s to 1.8 m/s. The simulation continues for 1 hour. Thirteen radioactive sources are placed at the intersection of roads[1]. The results of global and regional anomaly detection methods are shown in Fig. 4.2.



(a) Global anomaly score distribution.

(b) Regional anomaly score distribution.

Figure 4.2: Comparison of global and regional anomaly detection methods. 13 radioactive sources are placed at the intersections of roads, denoted by yellow color.

In Fig. 4.2, high background region is in the lower left part and the center of yellow dots represent the location of radioactive sources. For the global method in Fig. 4.2a, the average background radiation level is 33.96 *cps* and the standard deviation is 6.8 *cps*. For the regional method, the local background radiation level is estimated using the measurements within 30 meters and the average background radiation changes spatially as shown in Fig. 3.5. When the background radiation is estimated using the data from the entire region, the background radiation estimation, not surprisingly, overestimates the background radiation in the low background radiation region but underestimates the background radiation in the high background radiation region. In other words, it increases the false positive rate in the high background radiation region but reduces the true positive rate in the low background radiation region. This is better reflected in Fig. 4.2a, where the anomaly score is much higher in the lower left region compared with Fig. 4.2b. On the other hand, the background radiation estimated using regional measurements can reflect the true background radiation

[1]The exact locations of radioactive sources have little influence on the conclusion. For more details about the simulation, see Appendix E.

level, thus leading to overall better performance.

To better quantify the difference, using the calculated anomaly score shown in Fig. 4.2a and Fig. 4.2b, by setting different threshold values, true positive rate and false positive rate can be calculated. The ROC curves of global and regional methods based on calculated true positive rate and false positive rate are shown in Fig. 4.3.

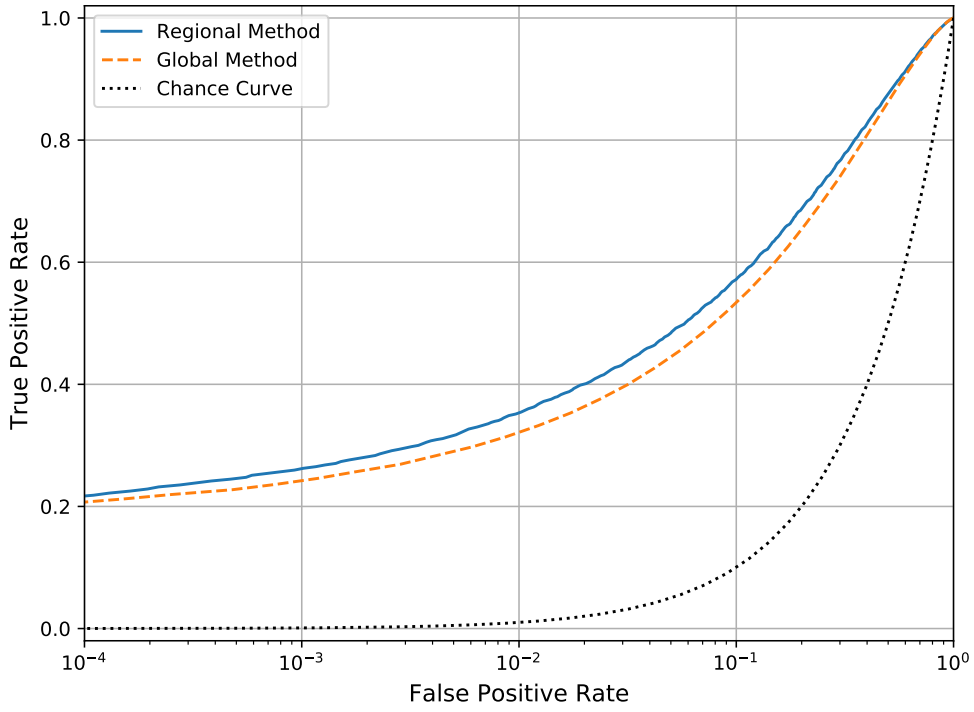


Figure 4.3: ROC curves of global and regional anomaly detection methods (regional method uses the data from surrounding 30 m radius to estimate the background radiation).

As shown in Fig. 4.3, the regional method is better than the global method, but the improvement is not large. This is because that the background radiation spatial variation, which is shown in Fig. 3.5, is not very large. When the spatial variation is larger, the improvement is expected to increase. In the global method, the background radiation only needs to be estimated once for the entire region. But for the regional method, this process needs to be repeated multiple times for the studied region, and each time, nearby measurements need to be queried from the database. These procedures will dramatically increase the computational complexity. In practical implementation, the background radiation spatial variation and computational complexity need to be considered together to make the appropriate choice between the regional and the global approaches.

4.3 Influence of Detector Characteristics

In addition to radiation detection algorithms, another factor that influences the performance of mobile sensor networks is the characteristic of radiation detectors. Different detectors may consist of different materials and have different sizes. These factors will affect the quality of the collected signals.

The influence of detector characteristics can be well quantified using the signal-to-noise ratio (SNR). In the nuclear radiation detection field, since the background radiation is generally considered as noise, the standard deviation of background radiation is used to represent the level of noise. The signal refers to the radiation from radioactive sources, which is simply the difference of measured signal and estimated background radiation. Thus, SNR [9] is defined as:

$$SNR(x) = \frac{s}{\sqrt{b}} = \frac{x-b}{\sqrt{b}}, \tag{4.3}$$

where s is the count rate from radioactive sources, x is the the measured count rate by the detector, and b represents the estimated background radiation.

There are different kinds of detectors. Here, two typical detectors, D3S and Ortec-905, are simulated using GADRAS. During the simulation, a ^{137}Cs source with source intensity of 500 μCi (626 cps) is used. The detectors are placed 1 m high above the ground. Simulated data are obtained at different distances from the radioactive source. SNRs are calculated using Eq. 4.3. The exact values for measured source count rate s and calculated the SNR using Eq. 4.3 at different distances are listed in Table 4.1.

Table 4.1: Source Intensity and SNR at Different Distances without Shielding

Configuration		**Distance to Radioactive Source**				
		1 m	**3 m**	**5 m**	**7 m**	**9 m**
D3S Detector	**s** (cps)	644.05	81.75	33.25	16.35	11.55
	SNR	116.64	14.64	6.12	3.10	2.11
Ortec 905 Detector	**s** (cps)	3369.25	605.85	255.85	174.50	91.60
	SNR	307.72	54.86	22.74	11.39	6.18

In the simulation, the 500 μCi (626 cps) ^{137}Cs source is placed in front of the detectors. There is no shielding materials between the source and the detectors.

The plot of SNR versus distance to the ^{137}Cs source is shown in Fig. 4.4. As the distance of the detector from the radioactive source increases, the SNR decreases dramatically for both detectors. And the SNR

for the Ortec 905 detector, not surprisingly, is much larger than that for the D3S detector with the same settings. However, better detectors are more expensive, and thus a trade-off between fewer more expensive detectors versus a large number of low cost detectors needs to be made.

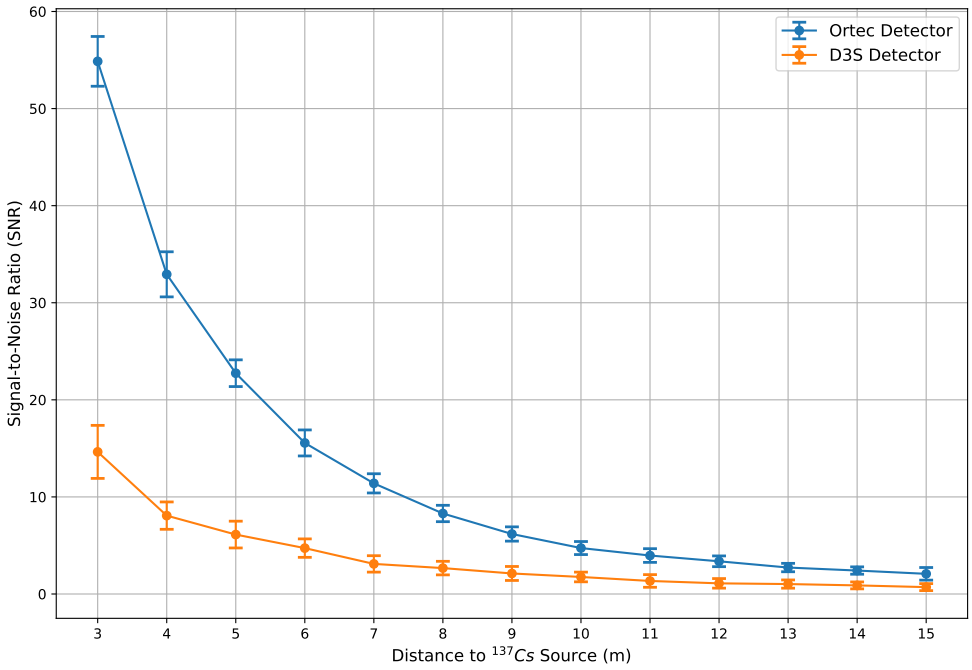


Figure 4.4: Comparing the signal-to-noise ratio (SNR) of Ortec 905-3 detector and D3S detector. The error bar is shown with one standard error.

4.4 Influence of Number of Detectors

Among all the influencing factors, the number of detectors[2] is certainly one important factor that determines the performance of mobile sensor networks. The number of detectors directly determines the probability of the radioactive source being detected within a limited time period. In other words, the number of detectors determines the time for the radioactive source to be detected, and finally determines the performance of the mobile sensor networks for nuclear radiation detection.

As more and more mobile detectors are deployed in the experimental region, the chance that the radioactive sources being detected is expected to increase. As a result, the time needed to detect the source decreases. In this work, the time needed to detect the source is defined as the time it takes one detector to reach within 10 meters of the source, or called reachable time. To better quantify the influence of the number of detectors, the simulation platform as described in Chapter 3 is used. The simulation is the same as

[2]It is more reasonable to use detector density instead of number of detectors for similar analyses. Here, since all simulations are conducted over the same area, the number of detectors can represent detector density.

described in Section 3.3, in which each detector is randomly moving with constant speed uniformly sampled from 1.4 m/s to 1.8 m/s. Thirteen simulated radioactive sources are placed at the intersection of roads[3]. The number of detectors is varied from 10 to 100, and the average reachable time is calculated for each case. The results are shown in Fig. 4.5.

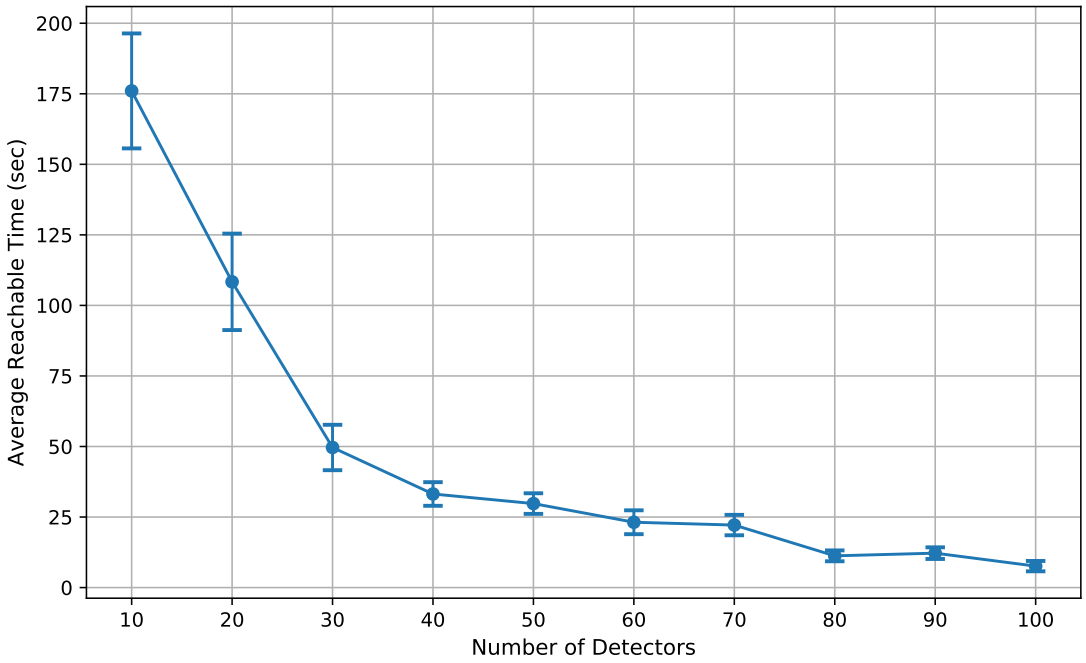


Figure 4.5: Average reachable time versus number of detectors. Error bar is shown with one standard error. The simulated map is shown in Fig. 3.4, which has the area of 650 m × 530 m.

As shown in Fig. 4.5, with more detectors, as expected, the average reachable time decreases. And as will be shown in Chapter 6, more detectors can reduce the time to locate the radioactive sources as well as increase the source localization accuracy. This is obviously to be expected. The goal of this exercise is to estimate the optimal number of detectors beyond which a marginal improvement in performance is not worth the additional cost[4]. It is clear that, as the number of detectors increases, the average reachable time decreases fast at first, then decreases slower and finally tends to saturate after the number of detectors is large enough. The optimal number of detectors for the current simulation setting (area covered and sources) seems to be between 30 and 40 considering the trade-off between cost and performance[5].

[3]For more details about the simulation, see Appendix E.

[4]Currently, a D3S detector costs about $5,800.

[5]Number of detectors is not the only factor that influences the reachable time. Another factor is the detector's moving speed. When detectors are moving slowly, it is expected to take a longer time for the radioactive source to be detected. However, when the speed is increased, although the reachable time is decreased, the quality of the collected data may get worse. This is because that when detector moves faster, the measured signal will contain more information from background since it will travel a longer distance.

To see the influence of number of detectors on the performance of mobile sensor networks for anomaly detection, under the same simulation settings, the number of detectors is again varied over 10 to 100 and the area under the curve (AUC) with the anomaly detection technique as described in Section 4.2 is calculated. The results are shown in Fig. 4.6.

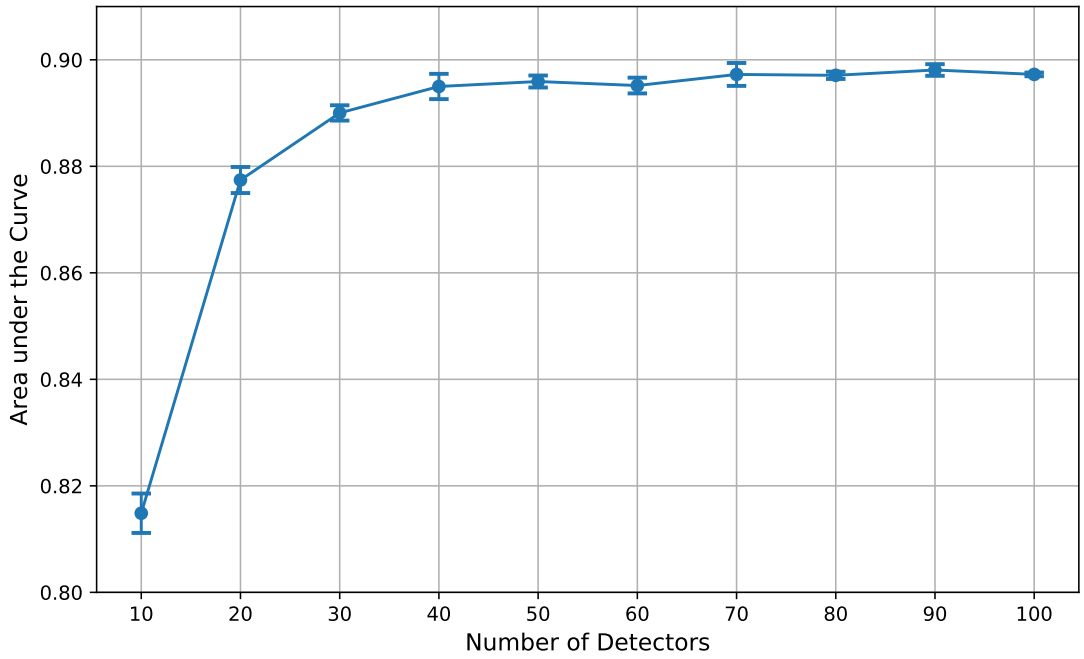


Figure 4.6: Area under the ROC curve with different number of detectors. The error bar is shown with one standard error (Due to the randomness in simulation, there are some fluctuations with the error bars).

As shown in Fig. 4.6, in a limited amount of time, increasing the number of mobile detectors as expected improves the performance of mobile sensor networks for anomaly detection, especially when the number of detectors is less than 40. However, the improvement diminishes when the number of detectors is large enough, which is consistent with the results shown in Fig. 4.5. Figure 4.6 again suggests that the optimal number of detectors to be between 30 and 40. The reason behind the saturation when detector number is larger than 40 is that current anomaly detection is based on single measurement from each detector. Once the estimated background radiation level is fixed, increasing the number of detectors will not change the Poisson anomaly score of each measurement.

4.5 Summary

Analysis of the performance of mobile sensor networks under different settings is presented in this chapter. Focused on anomaly detection, the influence of anomaly detection algorithms, detector characteristics, and the number of detectors are studied in detail. As expected, regional detection algorithms, better radiation detectors, and more detectors can help improve the mobile sensor network's performance for nuclear radiation detection. But this improvement may become negligible with increasing quality and quantity of radiation detectors. The trade-off between cost and performance is always important when designing mobile sensor networks for radiation detection. Through simulation, this chapter quantifies the influence of each factor, and can provide guidance for the actual deployment of mobile sensor networks.

Chapter 5

Poisson Kriging for Background Radiation Estimation and Anomalous Source Detection

5.1 Introduction

The influence of different parameters on the performance of mobile sensor networks were reported in Chapter 4. The final goal of mobile sensor networks, however, is not only to distinguish the abnormal measurements from normal background measurements, but also to locate the radioactive sources. The method proposed in Section 4.2 works when the background radiation is available. But there are cases in which mobile sensor networks can only cover part of the physical area, and background radiation measurements are not available for the rest of regions. In this case, it is necessary to develop methods to estimate the background radiation levels for regions where background radiation data are not available using measured data. In this chapter, Poisson kriging, which is based on geospatial modeling, is implemented to estimate the background radiation levels as well as to detect the anomalous radioactive sources.

The first law of geography [80] states that "*everything is related to everything else, but near things are more related than distant things*". This law also works for nuclear radiation, especially background radiation. Nuclear radiation measures the radiation levels at specific locations and times, which are influenced by many factors. Given that radiation measurements are usually discrete in space and time, it is important to estimate the background radiation levels for unmeasured locations using the measurements from surrounding locations.

Kriging is a geostatistical interpolation method that uses the given measurements to interpolate data at positions where data were not collected [81]. There are several commonly used kriging models, including ordinary kriging, simple kriging, and universal kriging. For radiation detection, all measurements are integers, and the signals are assumed to follow the Poisson distribution. In this case, the often-used Gaussian assumption is not suitable for modeling radiation data since the radiation might have relatively low mean values and the measurements are discrete. Also, for radiation data, goal is to estimate the distribution of nuclear radiation levels, which can be represented as the average radiation count rate, using the meaured

radiation count rate data. Therefore, it is more reasonable to use the Poisson kriging approach[1]. Monestiez et al. [82] proposed Poisson kriging to model the spatial distribution of Balaenoptera physalus (a kind of whale) using the sparse count data. McShane et al. [83] developed a similar model to analyze the spatially correlated count data. Recently, Bellier et al. [84] extended Poisson kriging to nonstationary hierarchical model for count data. Since then, Poisson kriging has been implemented in different areas, including cholera and dysentery incidence [85], cancer [86, 87], and wildlife population [82, 84].

This chapter is focused on the application of Poisson kriging method for estimating nuclear radiation distribution using mobile sensor networks. More precisely, section 5.2 focuses on the theory behind Poisson kriging, and its implementation for background radiation estimation. Section 5.3 discusses anomalous source detection problem using Poisson kriging. The performance of the proposed algorithm is demonstrated using the measured data in which simulated radioactive sources are injected.

5.2 Poisson Kriging and Background Radiation Estimation

5.2.1 Poisson Kriging

The raw nuclear radiation data measured using various radiation detectors are in integer format, which are assumed to follow the Poisson distribution [58]. Poisson kriging was originally developed based on the population weighted semi-variogram estimators and used to analyze the count data [82]. Compared with other kriging methods, Poisson kriging assumes that measured data follow Poisson distribution. The probability of a detector collects m counts in a unit time with expectation λ is expressed as:

$$p(x = m) = \frac{\lambda^m}{m!} \cdot e^{-\lambda}, \tag{5.1}$$

where λ represents the average radiation count rate at given position.

For radiation measurements, let random observation $X_{\boldsymbol{r}}$ be the count rate measurement at location $\boldsymbol{r}$. There exists some underlying distribution $Y_{\boldsymbol{r}}$ which represents the expected value of $X_{\boldsymbol{r}}$. In other words, $Y_{\boldsymbol{r}}$ represents λ in Eq. 5.1 at position $\boldsymbol{r}$. The goal of Poisson kriging is to estimate the distribution of Y using measured X. Given latent variable $Y_{\boldsymbol{r}}$, measured radiation count rate $X_{\boldsymbol{r}}$ is assumed to follow the Poisson distribution:

$$X_{\boldsymbol{r}}|Y_{\boldsymbol{r}} \sim Poisson(Y_{\boldsymbol{r}}). \tag{5.2}$$

[1]More specifically, compared with tranditional kriging or Gaussian process, Poisson kriging explicitly assumes that measured data follow Poisson distribution and it is used to model discrete integer values. Also, Poisson kriging aims to model the hidden variable, instead of directly modeling the measured radiation count rate.

Further, $Y_{\boldsymbol{r}}$ is assumed to be a positive random field with mean $\mu_{\boldsymbol{r}}$ and variance $\sigma^2_{\boldsymbol{r}}$ [82]. To simplify the problem, $\mu_{\boldsymbol{r}}$ and $\sigma^2_{\boldsymbol{r}}$ are assumed to be constant [88] [2]. Then it follows that:

$$E[Y_{\boldsymbol{r}}] = \mu \tag{5.3}$$

$$E[Y^2_{\boldsymbol{r}}] = \mu^2 + \sigma^2. \tag{5.4}$$

Based on the above assumptions, the conditional expectation and variance of $X_{\boldsymbol{r}}$, given expectation $Y_{\boldsymbol{r}}$, can be calculated as follows [82]:

$$E[X_{\boldsymbol{r}}|Y_{\boldsymbol{r}}] = Y_{\boldsymbol{r}} \tag{5.5}$$

$$Var[X_{\boldsymbol{r}}|Y_{\boldsymbol{r}}] = Y_{\boldsymbol{r}} \tag{5.6}$$

$$E[X^2_{\boldsymbol{r}}|Y_{\boldsymbol{r}}] = Y_{\boldsymbol{r}} + Y^2_{\boldsymbol{r}} \tag{5.7}$$

$$E[X_{\boldsymbol{r}}] = E[E[X_{\boldsymbol{r}}|Y_{\boldsymbol{r}}]] = \mu \tag{5.8}$$

$$Var[X_{\boldsymbol{r}}] = E[E[X^2_{\boldsymbol{r}}|Y_{\boldsymbol{r}}]] - E^2[X_{\boldsymbol{r}}] = \mu + \sigma^2 \tag{5.9}$$

$$E[X_{\boldsymbol{r}}X_{\boldsymbol{r}'}|Y] = Cov(X_{\boldsymbol{r}}, X_{\boldsymbol{r}'}|Y) + E[X_{\boldsymbol{r}}|Y_{\boldsymbol{r}}]E[X_{\boldsymbol{r}'}|Y_{\boldsymbol{r}'}] = \delta_{\boldsymbol{r},\boldsymbol{r}'}Y_{\boldsymbol{r}} + Y_{\boldsymbol{r}}Y_{\boldsymbol{r}'}, \tag{5.10}$$

where $\delta_{\boldsymbol{r},\boldsymbol{r}'}$ is the delta function (1 if $\boldsymbol{r} = \boldsymbol{r}'$ and 0 otherwise).

For Poisson kriging, considering two locations $\boldsymbol{r}$ and $\boldsymbol{r}'$, the observed radiation count rate $X_{\boldsymbol{r}}$ is assumed to not interact with $X_{\boldsymbol{r}'}$ directly. Instead, the average radiation level $Y_{\boldsymbol{r}}$ are assumed to be connected with $Y_{\boldsymbol{r}'}$ through their covariance. The covariance function $C_Y(\boldsymbol{r}, \boldsymbol{r}') = Cov(Y_{\boldsymbol{r}}, Y_{\boldsymbol{r}'})$ for Y is assumed to depend only on the distance $||\boldsymbol{r} - \boldsymbol{r}'||$ between two locations:

$$C_Y(\boldsymbol{r}, \boldsymbol{r}') = E[(Y_{\boldsymbol{r}} - \mu)(Y_{\boldsymbol{r}'} - \mu)] = E[Y_{\boldsymbol{r}}Y_{\boldsymbol{r}'}] - \mu^2 \tag{5.11}$$

[2] For non-constant mean cases, the trend can be estimated first and the problem is still solvable. See reference [84] for more details.

$$
\begin{aligned}
E[(X_{\boldsymbol{r}} - X_{\boldsymbol{r}'})^2|Y] &= E[X_{\boldsymbol{r}}^2|Y_{\boldsymbol{r}}] - 2E[X_{\boldsymbol{r}}X_{\boldsymbol{r}'}|Y] + E[X_{\boldsymbol{r}'}^2|Y_{\boldsymbol{r}'}] \\
&= Y_{\boldsymbol{r}}^2 + Y_{\boldsymbol{r}} + Y_{\boldsymbol{r}'}^2 + Y_{\boldsymbol{r}'} - 2\delta_{\boldsymbol{r},\boldsymbol{r}'}Y_{\boldsymbol{r}} - 2Y_{\boldsymbol{r}}Y_{\boldsymbol{r}'} \\
&= Y_{\boldsymbol{r}} + Y_{\boldsymbol{r}'} - 2\delta_{\boldsymbol{r},\boldsymbol{r}'}Y_{\boldsymbol{r}} + (Y_{\boldsymbol{r}} - Y_{\boldsymbol{r}'})^2
\end{aligned} \tag{5.12}
$$

$$
\begin{aligned}
E[(X_{\boldsymbol{r}} - X_{\boldsymbol{r}'})^2] &= E[E[(X_{\boldsymbol{r}} - X_{\boldsymbol{r}'})^2|Y]] \\
&= E[Y_{\boldsymbol{r}}] + E[Y_{\boldsymbol{r}'}] - 2E[\delta_{\boldsymbol{r},\boldsymbol{r}'}Y_{\boldsymbol{r}}] + E[(Y_{\boldsymbol{r}} - Y_{\boldsymbol{r}'})^2] \\
&= 2\mu - 2\delta_{\boldsymbol{r},\boldsymbol{r}'}\mu + E[(Y_{\boldsymbol{r}} - Y_{\boldsymbol{r}'})^2].
\end{aligned} \tag{5.13}
$$

Traditionally, it is more common to use the semi-variogram instead of covariance to model the correlation. The semi-variogram $\gamma(\boldsymbol{r}, \boldsymbol{r}')$ was first defined by Matheron [89] as half of the average squared difference between points $\boldsymbol{r}$ and $\boldsymbol{r}'$:

$$
\gamma(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{2}Var(Z(\boldsymbol{r}) - Z(\boldsymbol{r}')), \tag{5.14}
$$

where $Z(\boldsymbol{r})$ is the measured value and $Var(\cdot)$ is the variance. Here, for the measured radiation count rate X, the semi-variogram functions for X can be further simplified as $\gamma_X(\boldsymbol{r}, \boldsymbol{r}') = \frac{1}{2}E[(X_{\boldsymbol{r}} - X_{\boldsymbol{r}'})^2]$ based on Eq. 5.14. From Eq. 5.13, the semi-variogram function for Y can be calculated as:

$$
\gamma_Y(\boldsymbol{r}, \boldsymbol{r}') = \gamma_X(\boldsymbol{r}, \boldsymbol{r}') + \delta_{\boldsymbol{r},\boldsymbol{r}'}\mu - \mu. \tag{5.15}
$$

From Eq. 5.11 and Eq. 5.15, the covariance function $C_Y(\boldsymbol{r}, \boldsymbol{r}')$ and semi-variogram function $\gamma_Y(\boldsymbol{r}, \boldsymbol{r}')$ are related through:

$$
C_Y(\boldsymbol{r}, \boldsymbol{r}') = \sigma^2 - \gamma_Y(\boldsymbol{r}, \boldsymbol{r}'). \tag{5.16}
$$

For radiation detection with multiple detectors, suppose that there are n measurements $X_1, X_2, \cdots, X_n$ from different locations $\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_n$, the estimate of Y_0 for the unknown position $\boldsymbol{r}_0$ is assumed to be the linear combination of available measurement $X_1, X_2, \cdots, X_n$:

$$
\hat{Y_0} = \sum_{i=1}^{n} \lambda_i X_i. \tag{5.17}
$$

The problem becomes of finding $\lambda_1, \lambda_2, \cdots, \lambda_n$ such that Eq. 5.17 works as the optimal estimator so that

the estimator $\hat{Y}_0$ is unbiased and the squared estimation error $E[(\hat{Y}_0 - Y_0)^2]$ is minimized. More specifically, to ensure that $\hat{Y}_0$ is an unbiased estimator of Y_0, the expectation of estimator Y_0 should be equal to μ [82]:

$$E[\hat{Y}_0] = \sum_{i=1}^{n} \lambda_i E[X_i] = \mu \sum_{i=1}^{n} \lambda_i = \mu, \tag{5.18}$$

which leads to

$$\sum_{i=1}^{n} \lambda_i = 1. \tag{5.19}$$

Further, $E[(\hat{Y}_0 - Y_0)^2]$ can be derived as:

$$\begin{aligned} E[(\hat{Y}_0 - Y_0)^2|Y] &= E[(\sum_{i=1}^{n} \lambda_i X_i - Y_0)^2|Y] \\ &= E[(\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j X_i X_j - 2Y_0 \sum_{i=1}^{n} \lambda_i X_i + Y_0^2)|Y] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j Y_i Y_j + \sum_{i=1}^{n} \lambda_i^2 Y_i - 2Y_0 \sum_{i=1}^{n} \lambda_i Y_i + Y_0^2 \end{aligned} \tag{5.20}$$

$$\begin{aligned} E[(\hat{Y}_0 - Y_0)^2] &= E[E[(\hat{Y}_0 - Y_0)^2|Y]] \\ &= E[\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j Y_i Y_j + \sum_{i=1}^{n} \lambda_i^2 Y_i - 2Y_0 \sum_{i=1}^{n} \lambda_i Y_i + Y_0^2] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C_Y(i,j) + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \mu^2 + \sum_{i=1}^{n} \lambda_i^2 \mu - 2\sum_{i=1}^{n} \lambda_i C_Y(i,0) - 2\sum_{i=1}^{n} \lambda_i \mu^2 + \mu^2 + \sigma^2. \end{aligned} \tag{5.21}$$

From the unbiased condition in Eq. 5.18 and Eq. 5.19, Eq. 5.21 can be simplified as:

$$E[(\hat{Y}_0 - Y_0)^2] = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C_Y(i,j) + \sum_{i=1}^{n} \lambda_i^2 \mu - 2\sum_{i=1}^{n} \lambda_i C_Y(i,0) + \sigma^2. \tag{5.22}$$

And the variance of the estimation $Var(\hat{Y}_0 - Y_0)$ is:

$$\begin{aligned} Var(\hat{Y}_0 - Y_0) &= E[(\hat{Y}_0 - Y_0)^2] - E^2[\hat{Y}_0 - Y_0] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C_Y(i,j) + \sum_{i=1}^{n} \lambda_i^2 \mu - 2\sum_{i=1}^{n} \lambda_i C_Y(i,0) + \sigma^2. \end{aligned} \tag{5.23}$$

To minimize $E[(\hat{Y}_0 - Y_0)^2]$, using the method of Lagrange multipliers, the following $n+1$ equations are derived:

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{5.24}$$

$$\sum_{j=1}^{n} \lambda_j C_Y(i,j) + \lambda_i \mu + \alpha = C_Y(i,0) \quad for\ i = 1, 2, \cdots, n, \tag{5.25}$$

where α is the Lagrange multiplier. The prediction variance can be calculated as:

$$Var(\hat{Y}_0 - Y_0) = \sigma^2 - \sum_{i=1}^{n} \lambda_i C_Y(i,0) - \alpha. \tag{5.26}$$

Equation 5.24 and Eq. 5.25 can be written in matrix format $A\boldsymbol{x} = \boldsymbol{b}$. A is a $(n+1)$ by $(n+1)$ matrix, and $\boldsymbol{x}$ and $\boldsymbol{b}$ are both column vectors with $(n+1)$ elements. To simplify the expression, C_{ij} is used to represent the $C_Y(i,j)$. More precisely,

$$A = \begin{pmatrix} C_{11}+\mu & C_{12} & C_{13} & \cdots & C_{1n} & 1 \\ C_{21} & C_{22}+\mu & C_{23} & \cdots & C_{2n} & 1 \\ C_{31} & C_{32} & C_{33}+\mu & \cdots & C_{3n} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{n1} & C_{n2} & C_{n3} & \cdots & C_{nn}+\mu & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \tag{5.27}$$

$$\boldsymbol{x} = (\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_n, \alpha)^T \tag{5.28}$$

$$\boldsymbol{b} = (C_{10}, C_{20}, C_{30}, \cdots, C_{n0}, 1)^T. \tag{5.29}$$

5.2.2 Implementation

The previous section describes the mathematics behind the Poisson kriging model, which leads to solving $A\boldsymbol{x} = \boldsymbol{b}$. From Eq. 5.27, Eq. 5.28, and Eq. 5.29, the key to solving the Poisson kriging problem is to calculate the covariance function C_{ij}. Since $C_{ij} = C_Y(i,j) = \sigma^2 - \gamma_Y(i,j)$, to estimate C_{ij}, the semi-variogram function $\gamma_Y(i,j)$ needs to be estimated first. Semi-variogram function $\gamma_Y(i,j)$ is assumed to be determined by the distance $d_{ij} = ||\boldsymbol{r}_i - \boldsymbol{r}_j||$ only, which is denoted by $\hat{\gamma}_Y(d_{ij})$.

Based on Eq. 5.15, the semi-variogram function of Y can be calculated from the semi-variogram of X, and $\gamma_X(h)$ is generally estimated [82] using:

$$\hat{\gamma}_X(h) = \frac{1}{N(h)} \sum_{i,j} \left[\frac{1}{2} \left(X_i - X_j \right)^2 \right]_{I_{d_{ij} \approx h}}, \tag{5.30}$$

where $N(h) = \sum_{i,j} I_{d_{ij} \approx h}$. $I_{d_{ij} \approx h}$ is the indicator function which is 1 when the distance between $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ is roughly equal to h. Otherwise, this indicator function is equal to 0. From Eq. 5.15, the semi-variogram function $\gamma_Y(h)$ can be easily inferred.

Equation 5.30 only gives discrete estimates of $\hat{\gamma}_X(h)$. To get a continuous estimate, a semi-variogram model can be fitted based on the calculated $\hat{\gamma}_X(h)$ for discrete distance h. Commonly used semi-variogram models [81] include:

Nugget model:
$$g(h) = \begin{cases} c_0 & \text{if } h = 0 \\ c_0 + c & \text{otherwise} \end{cases} \tag{5.31}$$

Spherical model:
$$g(h) = \begin{cases} c_0 + c \cdot (1.5 \cdot (\frac{h}{a}) - 0.5 \cdot (\frac{h}{a})^3) & \text{if } h \leq a \\ c_0 + c & \text{otherwise} \end{cases} \tag{5.32}$$

Exponential model:
$$g(h) = c_0 + c \left(1 - exp \left(-\frac{h}{a} \right) \right) \tag{5.33}$$

Gaussian model:
$$g(h) = c_0 + c \left(1 - exp \left(-\frac{h^2}{a^2} \right) \right) \tag{5.34}$$

Power model:
$$g(h) = c_0 + c \cdot h^{\omega} \quad \text{with} \quad 0 < \omega < 2. \tag{5.35}$$

In this work, the exponential model $\hat{\gamma}_X(h) = c_0 + c \left(1 - exp \left(-\frac{h}{a} \right) \right)$ is used[3]. The parameters c_0, c, and a are determined using the measured data.

5.2.3 Background Radiation Estimation

Previous sections described the mathematics behind the Poisson kriging model and necessary procedures to implement the model. To test the performance of the Poisson kriging model for background radiation

[3]Spherical, exponential, and Gaussian models have similar shapes, and the results are quite similar. The fitting model can be chosen based on minimizing the mean squared error of the fitted semi-variogram model or from previous research work [82].

estimation, background radiation data were measured using D3S detectors. The collected background radiation measurements described in section 3.4 are shown in Fig. 5.1, in which high background radiation regions are around the buildings, such as the statue and the church.

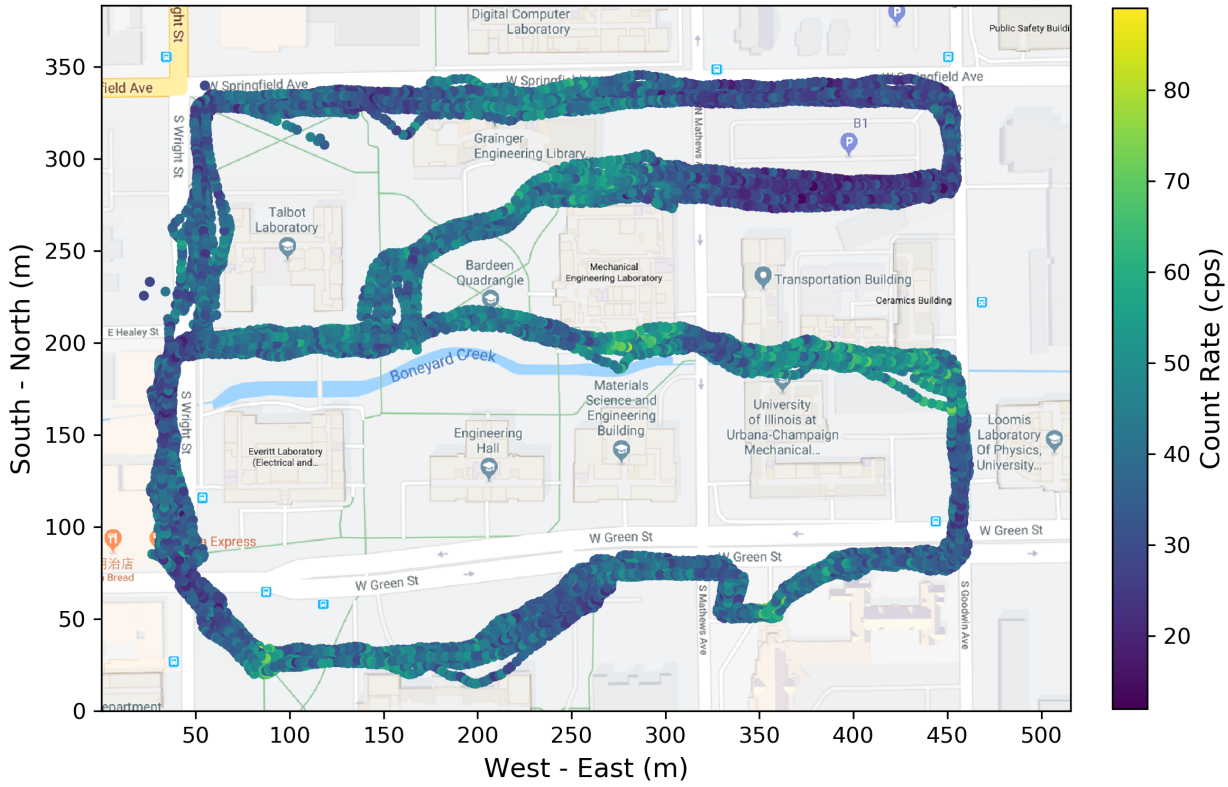


Figure 5.1: Background radiation measurements. Data were collected using D3S detectors in the outdoor environment.

The data collection was continued through the entire day. Using the background radiation data measured in the morning, discrete values for semi-variogram are calculated using Eq. 5.30. The calculated values and the fitted exponential semi-variogram model are shown in Fig. 5.2. The figure clearly shows that, for small distances, as the distance increases, the semi-variogram value increases. Since semi-variogram is negatively correlated with covariance, this trend is consistent with the expectation that radiation measurements are less correlated as distance increases. As the distance keeps increasing, the increase of semi-variogram diminishes and finally it tends to saturate. This is also consistent with the expectation that as the distance becomes large enough, there is no correlation between the two locations.

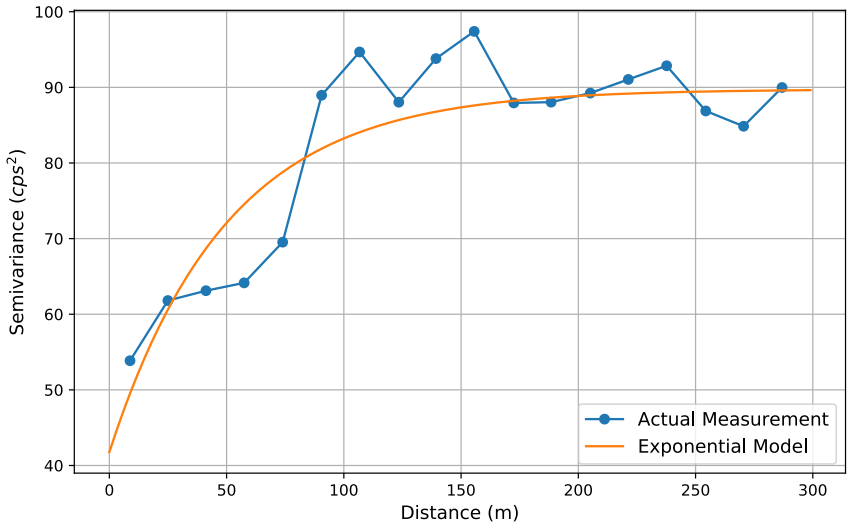


Figure 5.2: Semi-variogram plot of background radiation measurements.

Based on the estimated semi-variogram function shown in Fig. 5.2 and the background radiation measurements shown in Fig. 5.1, the continuous distribution estimation of background radiation levels is shown in Fig. 5.3. Compared with Fig. 5.1, Fig. 5.3 provides the smooth distribution of background radiation, and the high background radiation regions near the buildings can be easily identified.

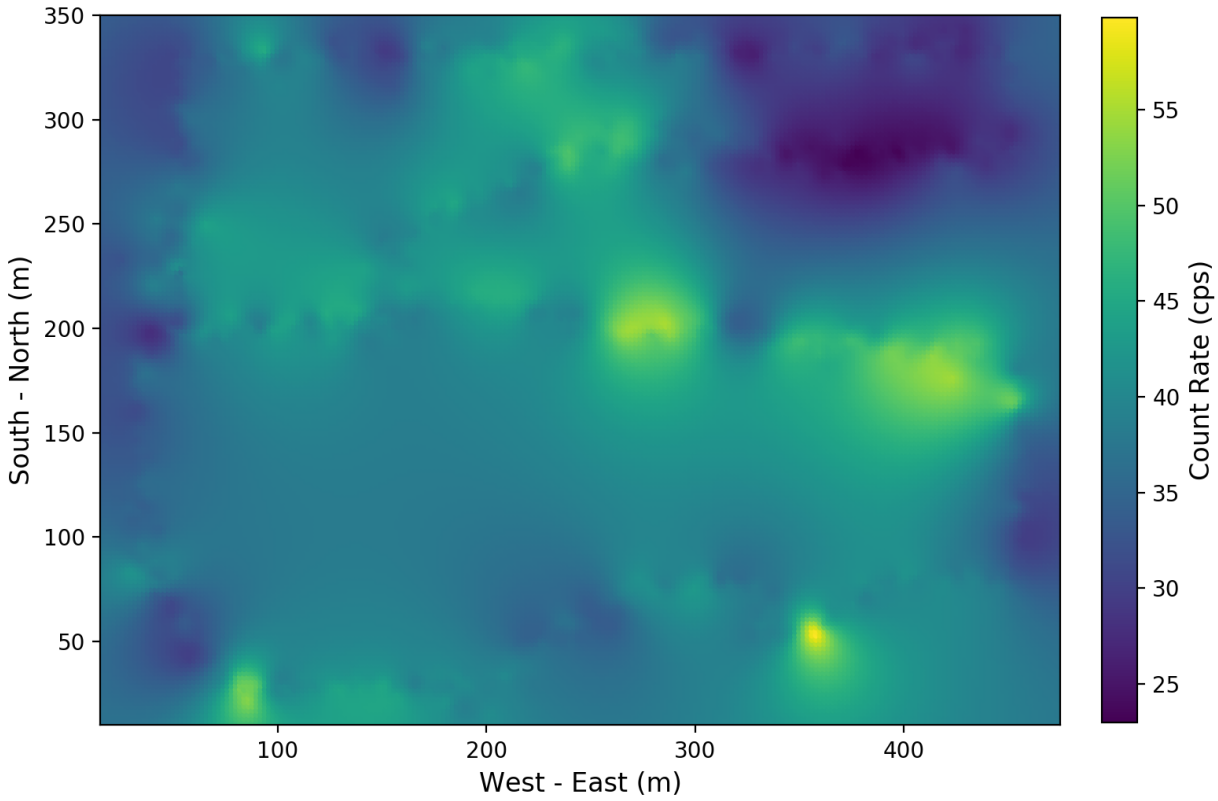


Figure 5.3: Poisson kriging estimation of background radiation distribution.

Also, based on Eq. 5.26, the standard deviation of the estimation can be easily calculated. The standard deviation corresponding to Fig. 5.3 is show in Fig. 5.4. As expected, there is less uncertainty for the locations that are close to radiation measurements. For the locations that are far away from any radiation measurements, the uncertainty is much larger.

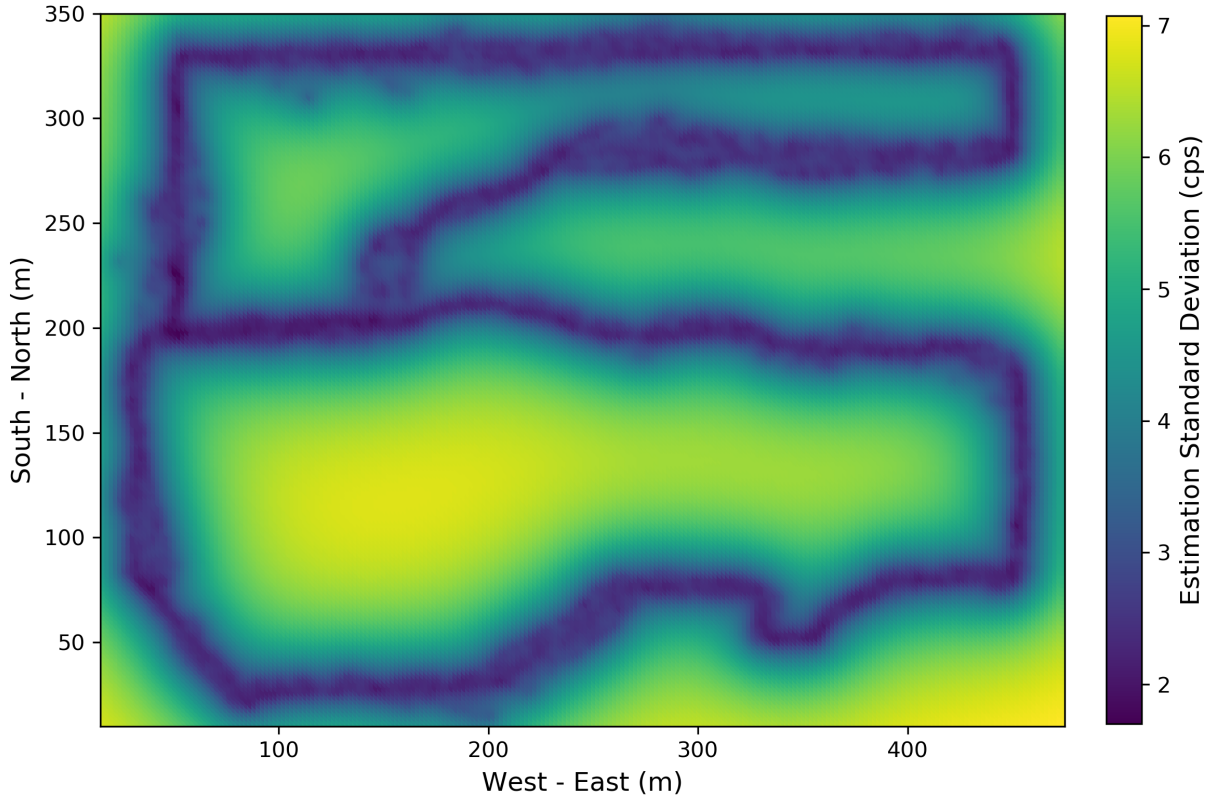


Figure 5.4: Distribution of standard deviation for Poisson kriging estimation.

5.3 Poisson Kriging for Anomalous Source Detection

5.3.1 Anomalous Source Detection

The details of Poisson kriging model and its application to background radiation estimation were covered in section 5.2. Background radiation estimation is usually the first step for nuclear radiation detection. Given the radiation data collected from mobile sensor networks, the ultimate goal is to detect and locate the radioactive sources. The application of Poisson kriging for anomalous source detection is discussed in this section.

After the mobile sensor network is established, normal background radiation data can be easily obtained since radiation detectors continuously collect radiation data in the experimental region covered by the mobile sensor network. Considering the fact that incidents involving nuclear radiation have extremely low probability, most of the data should contain no information about radioactive sources. Using the measured background radiation data, the background radiation distribution can be obtained using the Poisson kriging model as described in the previous section. This distribution will then become the base against which new data will be compared for source detection and localization. Once new measurements are obtained, different methods can then be applied to detect the anomalous radioactive sources.

For the experimental data shown in Fig. 5.1, the data collected in the morning is used to estimate the background radiation levels using Poisson kriging. For the data collected in the afternoon, synthetic radioactive sources are injected with different source intensity and locations, as described in section 3.4. Figure 5.5 shows the new measurements after source injection. In this sample case, a 100 μCi ^{137}Cs radioactive source is injected 1 meter away from the walking path, at the location denoted by a red cross inside the circle.

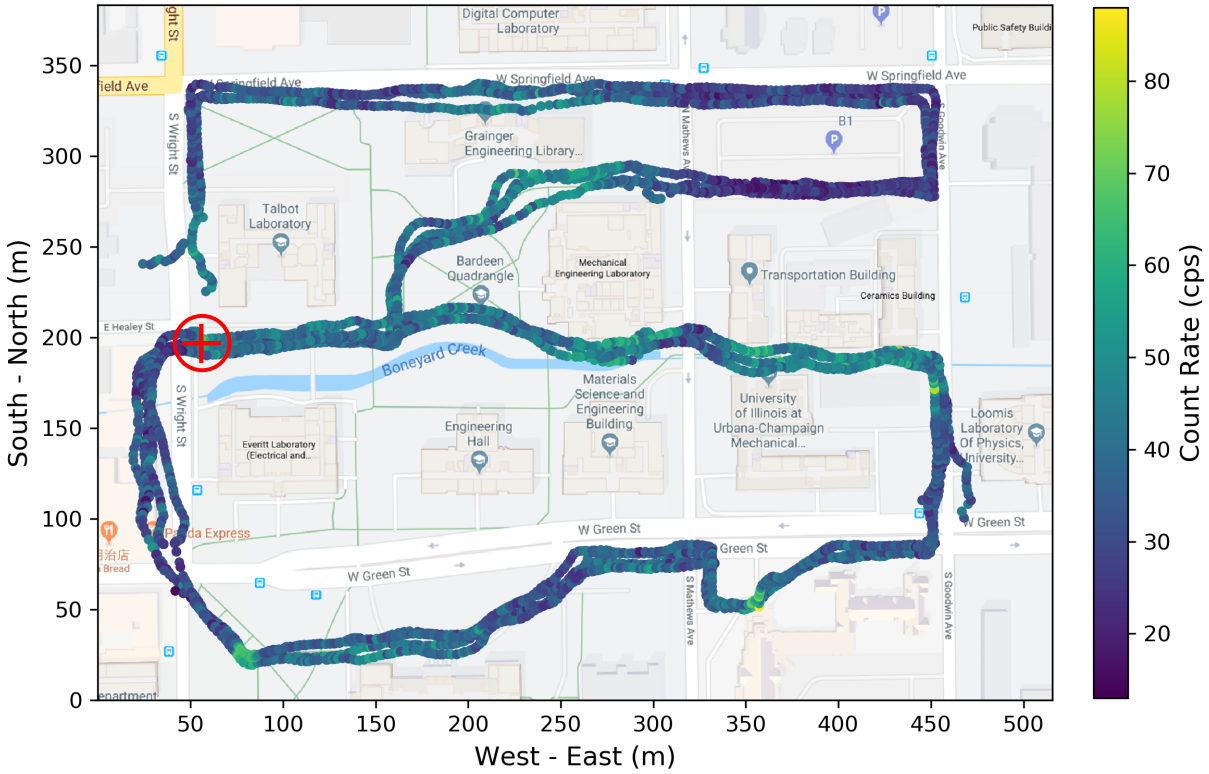


Figure 5.5: Radiation measurements with synthetic source injection. The injected source is denoted by the red cross inside the circle (the exact location is: x=55.87 m, y=196.93 m).

Point Data Based Method

After obtaining the estimation of background radiation levels as shown in Fig. 5.3 and new measurements (possibly with injected source data) as shown in Fig. 5.5, one method for source detection is based on point calculation. Similar to the procedures proposed by Reinhart [88], for the given position $\boldsymbol{r}$, suppose that the new measurement is denoted by $X^n_{\boldsymbol{r}}$, the estimate of background radiation level from Poisson kriging at position $\boldsymbol{r}$ is $Y^b_{\boldsymbol{r}}$, and $P(X^n_{\boldsymbol{r}}; Y^b_{\boldsymbol{r}})$ denotes the probability of measuring $X^n_{\boldsymbol{r}}$ when the expectation of Poisson distribution is $Y^b_{\boldsymbol{r}}$. Then, the Poisson anomaly score proposed in Eq. 4.2 can be used to detect radioactive sources:

$$AS_p(X^n_{\boldsymbol{r}}; Y^b_{\boldsymbol{r}}) = \begin{cases} \frac{1}{2}\frac{P(X^n_{\boldsymbol{r}}; Y^b_{\boldsymbol{r}})}{P(\lfloor Y^b_{\boldsymbol{r}} \rceil; Y^b_{\boldsymbol{r}})} & \text{if } X^n_{\boldsymbol{r}} \leq Y^b_{\boldsymbol{r}} \\ 1 - \frac{1}{2}\frac{P(X^n_{\boldsymbol{r}}; Y^b_{\boldsymbol{r}})}{P(\lfloor Y^b_{\boldsymbol{r}} \rceil; Y^b_{\boldsymbol{r}})} & \text{if } X^n_{\boldsymbol{r}} > Y^b_{\boldsymbol{r}} \end{cases}. \tag{5.36}$$

Equation 5.36 gives an anomaly score based on Poisson distribution in the range of $[0, 1]$. A higher anomaly score indicates a higher probability of containing information from radioactive sources. To test its performance in real cases, the experimental data with simulated source data injected as shown in Fig. 5.5 are used. Using this method, the distribution of the calculated anomaly score is shown in Fig. 5.6. Source detection can be carried out using the Poisson score as defined in Eq. 5.36.

In Fig. 5.6, the Poisson anomaly score is calculated only at the locations where new measurements are collected in the afternoon. This is because Eq. 5.36 is based on point estimation. Although this method can identify a collection of suspicious measurements around the injected radioactive sources, it also leads to many false positives at other locations due to the randomness of Poisson distribution. The ROC curve based on the results shown in Fig. 5.6 is shown in Fig. 5.7. The AUC is only 0.542, which means it is hard to confidently locate the radioactive sources.

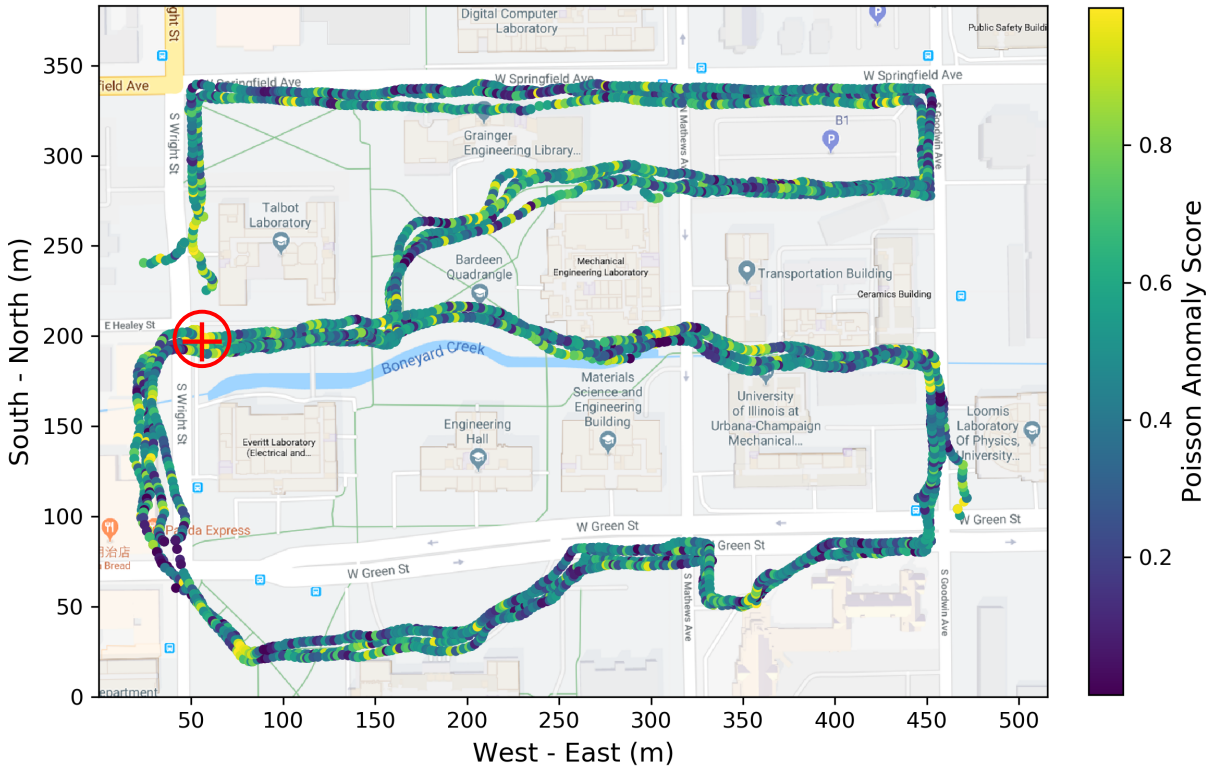


Figure 5.6: Poisson anomaly score distribution calculated using Poisson kriging model for background radiation. The injected source is denoted by the red cross inside the circle.

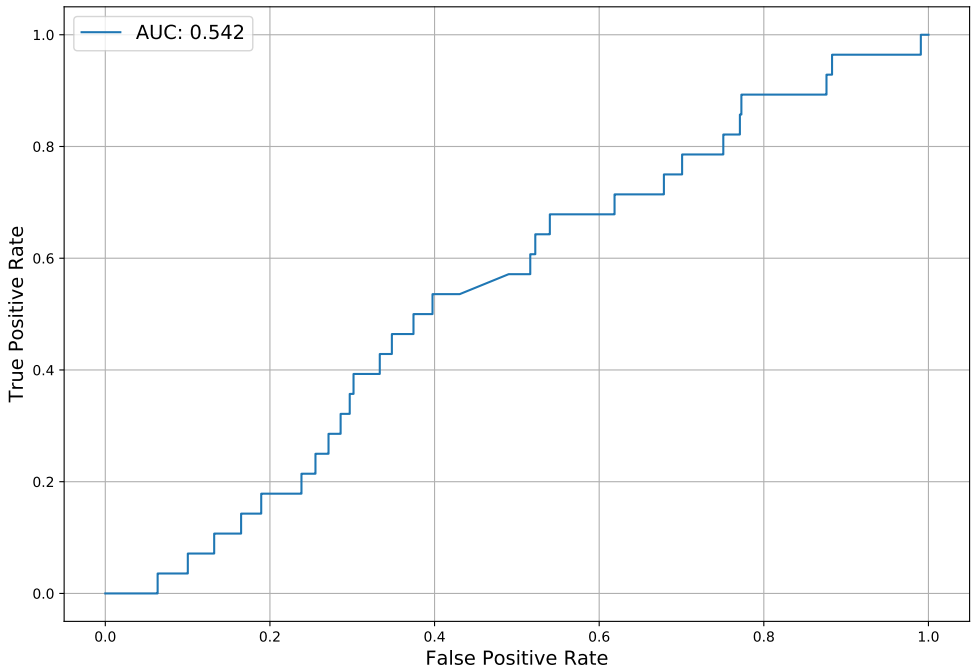


Figure 5.7: ROC curve for point data based method.

Distribution-Based Method

To overcome these difficulties, instead of comparing point values, a new anomaly detection technique based on comparing the new estimation against the original estimation is proposed. Suppose that the background radiation estimation for position $\boldsymbol{r}$ is denoted by $Y_{\boldsymbol{r}}^b$. When new data are collected in the experimental area, the estimation of current radiation level $Y_{\boldsymbol{r}}^n$ can be obtained through Poisson kriging carried out using the new measurements[4]. For the data shown in Fig. 5.5, radiation distribution after the source is injected is obtained using Poisson kriging model, and the result is shown in Fig. 5.8.

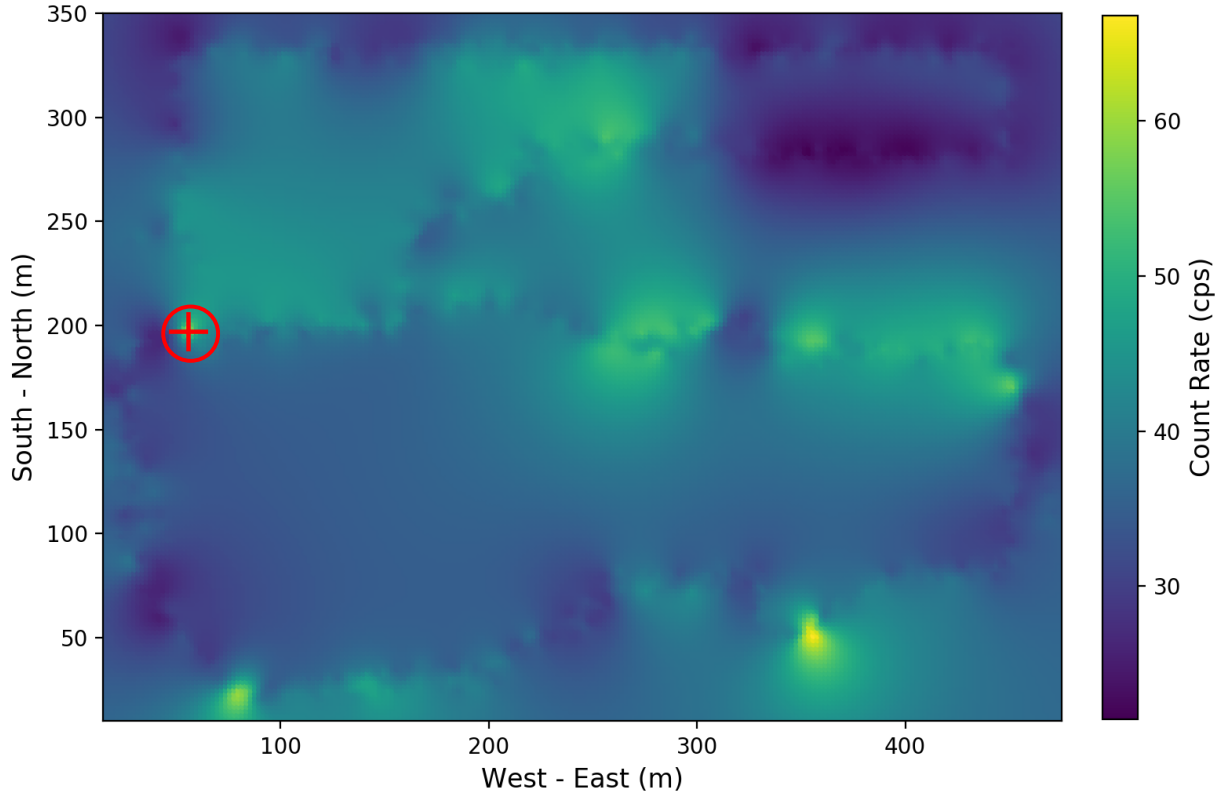


Figure 5.8: Poisson kriging estimation with synthetic source injection. The injected source is denoted by the red cross inside the circle.

In Fig. 5.8, compared with Fig. 5.3, the radiation level is clearly higher in the region surrounding the injected source. Thus, the difference between the new estimation $Y_{\boldsymbol{r}}^n$ and the estimation of background radiation estimation $Y_{\boldsymbol{r}}^b$ can be used to define an anomaly score:

$$Score(Y_{\boldsymbol{r}}^n; Y_{\boldsymbol{r}}^b) = Y_{\boldsymbol{r}}^n - Y_{\boldsymbol{r}}^b. \tag{5.37}$$

[4]In the cases that do not have enough measurements to build new semi-variogram model, the original semi-variogram model built using the background radiation measurements can be used.

Equation 5.37 simply compares the value of the new estimation with the estimation of background radiation level. It works when the background radiation distribution is relatively smooth. However, high background and low background radiation regions always exist, and due to the existence of high background radiation regions, the radioactive source cannot be easily detected. In fact, one of the spots in Fig. 5.8 with high background radiation shows a higher level of radiation than the region at the source.

SNR-Based Method

A more robust definition of anomaly score can be introduced by normalizing the score calculated in Eq. 5.37 by the background noise level. Using the square root of the estimated background radiation level as the estimation of noise level and the difference between new estimation and background radiation level as the estimation of signal strength, the anomaly score based on signal-to-noise ratio (SNR)[5] is defined as:

$$AS_{SNR}(Y_{\boldsymbol{r}}^n; Y_{\boldsymbol{r}}) = \frac{Y_{\boldsymbol{r}}^n - Y_{\boldsymbol{r}}^b}{\sqrt{Y_{\boldsymbol{r}}^b}}. \tag{5.38}$$

The anomalous source detection process corresponding to Eq. 5.38 is shown in Fig. 5.9. Once the mobile sensor network system is deployed, the background radiation measurements are available and the estimation of background radiation levels is obtained from Poisson kriging. After new measurements are collected, the new estimation of radiation distribution is obtained and the anomaly score based on SNR, $AS_{SNR}(Y_{\boldsymbol{r}}^n; Y_{\boldsymbol{r}})$, can be calculated for all locations. The location $\boldsymbol{r}$ with the highest anomaly score is assumed to be the most likely source location[6].

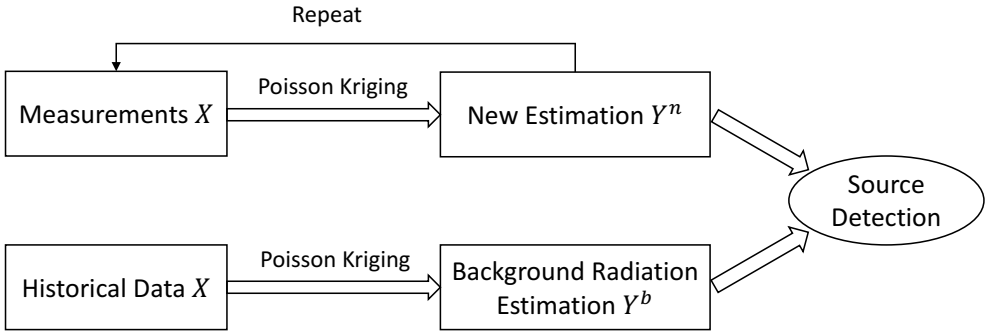


Figure 5.9: Anomalous source detection flowchart.

Using Eq. 5.38, the distribution of SNR-based anomaly score, AS_{SNR}, can be calculated in the studied

[5]Different from the Poisson anomaly score defined in Eq. 5.36, SNR based anomaly score is no longer bounded within the range of 0 to 1. There could be values larger than 1 as well as negative values.

[6]In most cases, the location with the highest anomaly score may not be the exact location of radioactive sources. In more practical implementations, certain threshold can be set and the regions that are above the given threshold are suspicious and should be identified for further exploration.

area, and the results are shown in Fig. 5.10. The actual source location is denoted by the red cross inside the circle and the estimated source location is denoted by the black triangle, which has the largest SNR. Clearly, compared with Fig. 5.8, the influence of the high background radiation region is mostly eliminated, and the radioactive source is correctly located.

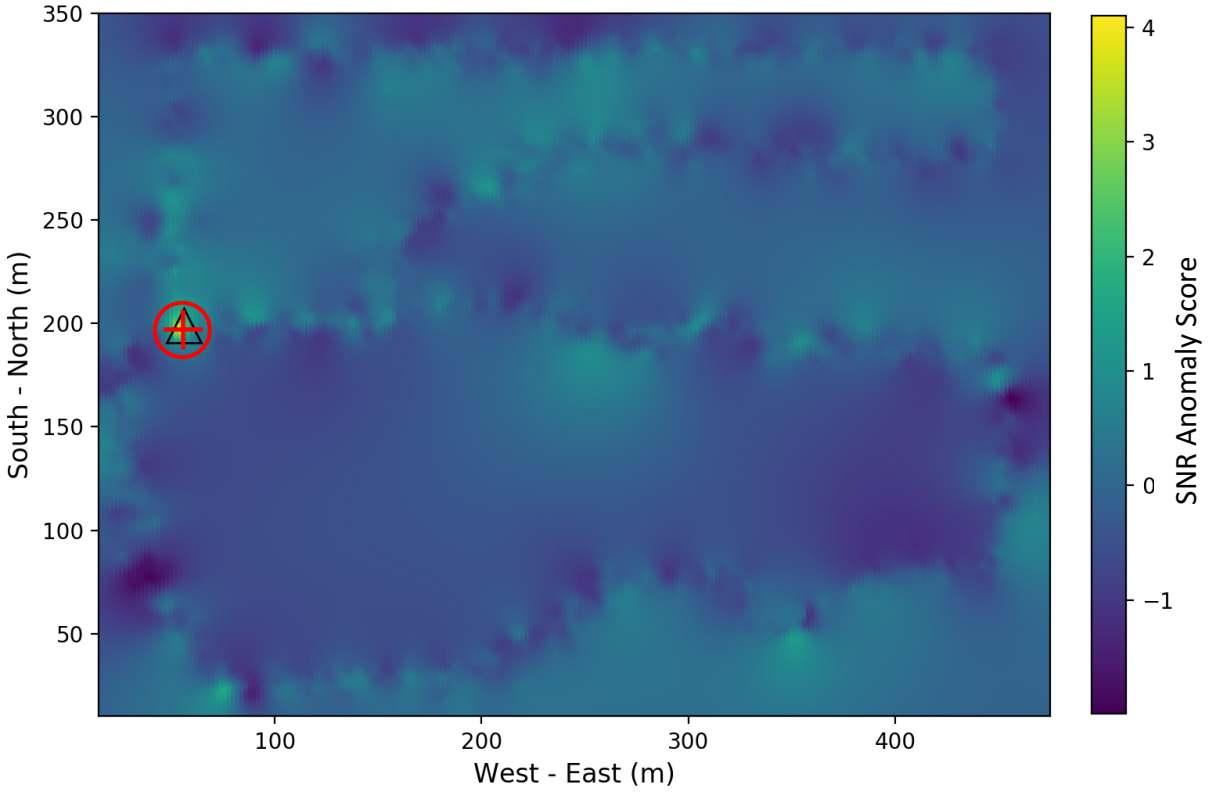


Figure 5.10: SNR-based anomaly score from the Poisson kriging model. Location of the injected source is denoted by the red cross inside a circle and the estimated source location is denoted by the black triangle.

5.3.2 Error Analysis

The previous section presented the application of Poisson kriging proposed in this chapter for anomalous source detection, and the results of a sample case where a 100 μCi source is injected 1 meter away from the walking paths are shown. In addition to the sample case, a comprehensive study of all the cases (source intensity varies from 100 μCi to 2,000 μCi and the distance to walking paths varies from 1 meter to 10 meters) shows that there are still some cases where the proposed algorithm could not correctly detect the radioactive sources. A sample case in which a 100 μCi ^{137}Cs source is injected 5 meters away from the walking path is shown in Fig. 5.11.

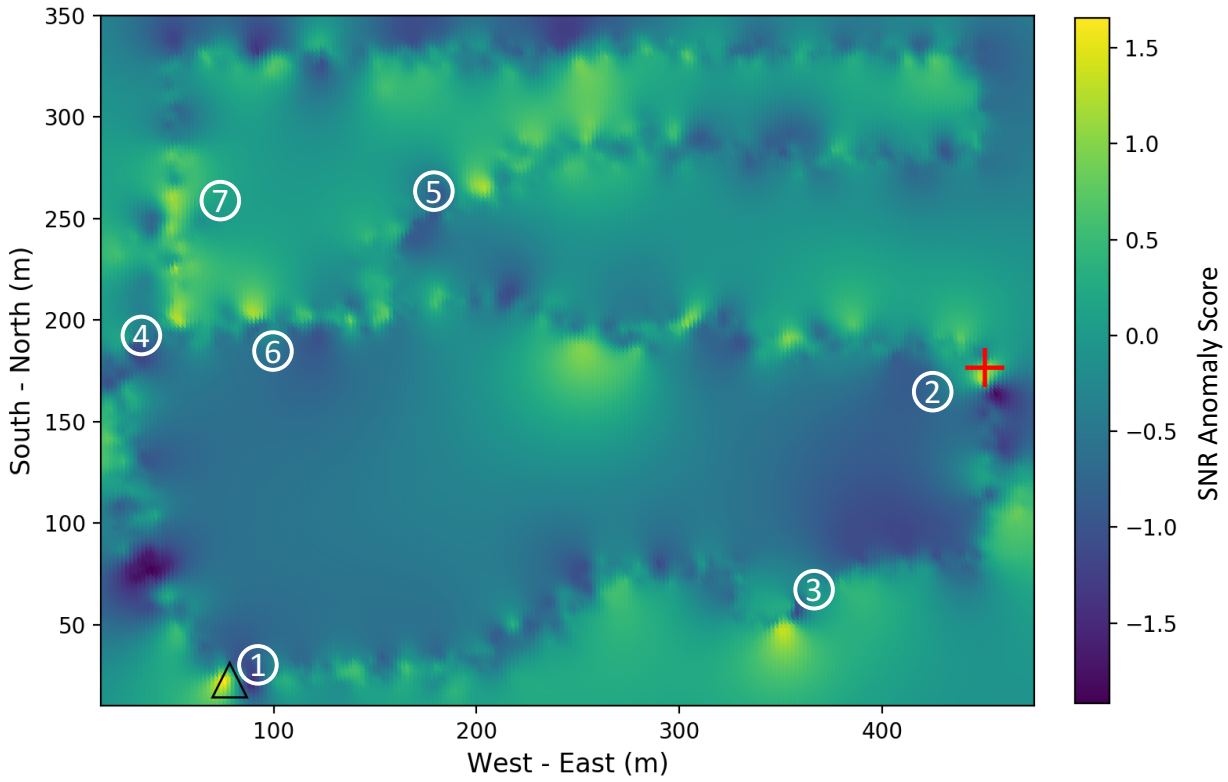


Figure 5.11: A case where the Poisson kriging model cannot detect correct radioactive source location. The injected source is denoted by a red cross and the estimated source location is denoted by a black triangle.

In Fig. 5.11, the injected source is denoted by the red cross and the estimated source location identified by the proposed method is around the statue (denoted by the dark triangle), one of the three locations with high background radiation level, instead of around the actual source location. One of the major reasons behind this is that the source intensity is low and the source is relatively far away from the walking path. The radiation level increase due to injected radioactive source is small. Another reason is that the measured background radiation data around the statue is higher in the afternoon compared with the measured radiation data in the morning[7]. In this way, the estimated radiation level around the statue is higher than the estimation obtained using the data collected in the morning, which leads to larger SNR-based anomaly score around the statue. For example, the SNR-based anomaly score at the actual source location is 1.38, while the anomaly score at the statue is 1.66. These two factors lead to the false positive detection by the proposed algorithm. If the simulated source is injected into the measured data in the morning, instead of the data in the afternoon, the proposed algorithm can still correctly locate the radioactive source. Refer to Appendix G

[7]Since in the morning and in the afternoon, all the conditions are quite similar except temperature, it is assumed that the high background radiation is partially due to the increase of temperature. For example, the higher temperature may cause change in the soil condition and cause higher background radiation. In addition, the radiation from cosmic rays may be another reason.

for additional details.

In the cases shown in Fig. 5.11, although the proposed method could not correctly locate the radioactive sources based on maximal SNR-based anomaly score, further study shows that the proposed method could still correctly identify the local hot-spots around the actual source location. For example, the proposed algorithm identifies 7 locations that have relatively larger anomaly scores. Ordered by the anomaly score in decreasing order, the top 7 potential locations are denoted by numbers from 1 to 7 on the figure, and the exact locations and corresponding SNRs are listed in Table 5.1. The largest SNR-based anomaly score happens around the statue (location 1), while the second largest SNR-based anomaly score is around the actual radioactive source (location 2). This kind of local hot-spot information is useful for restricting the suspicious region to smaller regions for further exploration.

Table 5.1: Top 7 Potential Radioactive Source Positions

No.	**x (m)**	**y (m)**	**SNR**
1	78.0	23.0	1.6574
2	452	173.0	1.3897
3	352.0	49.0	1.3596
4	56.0	201.0	1.2674
5	206.0	267.0	1.2293
6	92.0	205.0	1.2203
7	52.0	261.0	1.1944

In addition to missing the source as is the case shown in Fig. 5.11, another concern is the likelihood of the proposed method to predict false positives, most likely at high background regions. Figure 5.12 shows the results of the proposed methods for the measured data without any synthetic source being injected. The global hot-spot is denoted by the black triangle, which has the largest SNR-based anomaly score, and it is within the high background radiation region.

As a comparison, Fig. 5.13 shows the result of the same method on the data with a 100 μCi synthetic source injected 1 meter away from the walking path within the high background region. Figure 5.12 and Fig. 5.13 both identify the source to be within the high background region. However, compared with Fig. 5.12, in Fig. 5.13, the SNR-based anomaly score when actual radioactive source does exist is much larger than the cases without any radioactive source. For example, AS_{SNR} for the hotspot in Fig. 5.12 is 1.65 while it is 3.51 for the hotspot in Fig. 5.13. Thus, the false positive rate can be reduced by defining the appropriate threshold for the anomaly score.

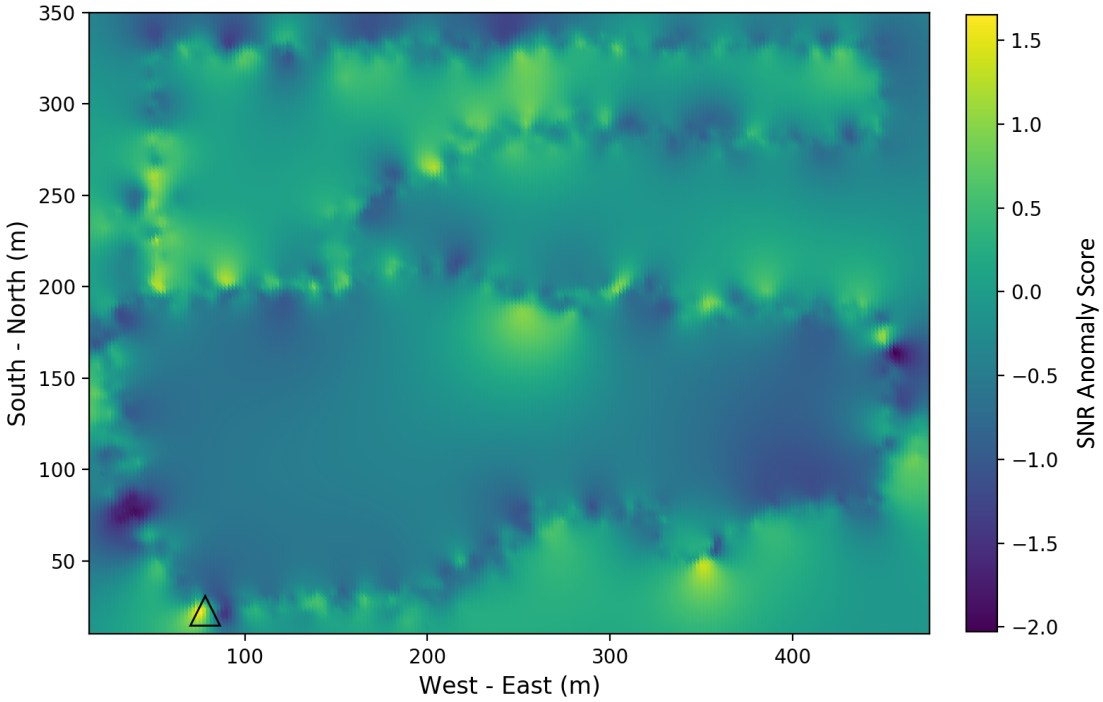


Figure 5.12: Anomaly score without radioactive sources. The estimated source location is denoted by the black triangle.

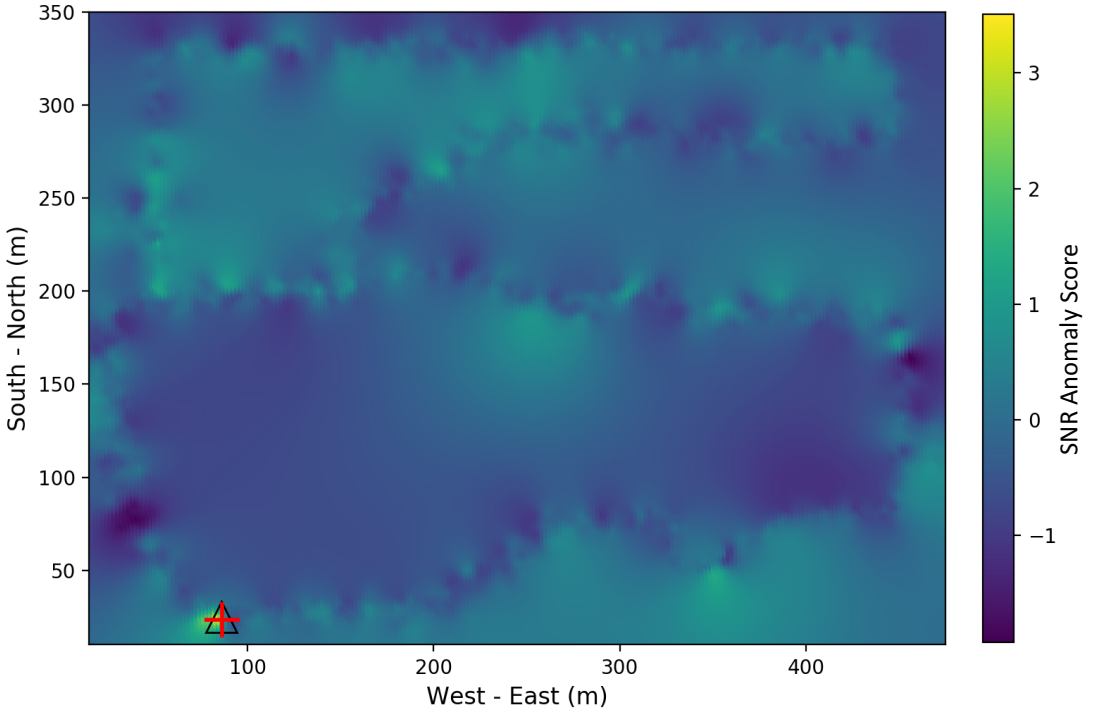


Figure 5.13: Anomaly score with source injected at high background region. The injected source is denoted by the red cross and the estimated source location is denoted by the black triangle.

To show the overall performance of the proposed methods for anomalous source detection problems, Table 5.2, Table 5.3, and Table 5.4 show the source detection accuracy with different distance thresholds[8]. The source detection accuracy is calculated according to the given distance threshold. When the distance of the estimated source location from the actual source location is smaller than the distance threshold, the source is assumed to be correctly detected.

As expected, the performance of the proposed method is dramatically influenced by source intensity and the distance of radioactive source from the walking paths. When the source is close enough or the source intensity is high enough or both, the detection accuracy could reach 100%. However, when the source intensity is low or the distance of radioactive source from the walking paths is too large, source cannot be easily detected based on global anomaly score[9]. But as discussed previously, when the proposed algorithm cannot correctly detect radioactive sources using the global maximal anomaly score, it can still identify local hot-spot regions around the radioactive sources as shown in Fig. 5.11.

Table 5.2: Source (^{137}Cs) Detection Accuracy with 5 m Distance Threshold

Configuration		**Distance to the Walking Path**		
		1 m	**5 m**	**10 m**
Source Intensity	100 μCi (127 cps)	80%	0%	0%
	300 μCi (380 cps)	100%	40%	0%
	500 μCi (626 cps)	100%	50%	20%
	1,000 μCi (1,220 cps)	100%	60%	30%
	2,000 μCi (2,354 cps)	100%	70%	30%

[8]It is worth clarifying the difference between distance threshold and distance to the walking path. Distance to the walking path is the distance of injected source from the walking path and it has the values of 1 meter, 5 meters and 10 meters. The distance threshold, on the other hand, is only used to calculate the source detection accuracy. In this work, distance thresholds are set to be 5 meters, 10 meters, and 20 meters for Table 5.2, Table 5.3, and Table 5.4, respectively.

[9]There are certain cases, for example, when the 100 μCi source is placed 10 meters away from the walking path, the 0% or 10% source detection accuracy is actually due to the fact that the nearby detectors cannot measure any useful signal. This is because that the source intensity is too low and the distance is too large. These cases can be totally ignored when evaluating the performance of the proposed algorithm.

Table 5.3: Source (^{137}Cs) Detection Accuracy with 10 m Distance Threshold

Configuration		**Distance to the Walking Path**		
		1 m	**5 m**	**10 m**
Source Intensity	100 μCi (127 cps)	90%	20%	10%
	300 μCi (380 cps)	100%	90%	20%
	500 μCi (626 cps)	100%	100%	40%
	1,000 μCi (1,220 cps)	100%	100%	50%
	2,000 μCi (2,354 cps)	100%	100%	60%

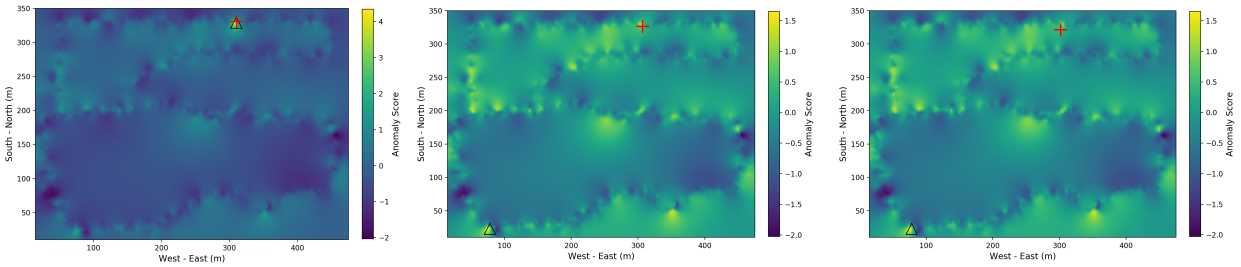
Table 5.4: Source (^{137}Cs) Detection Accuracy with 20 m Distance Threshold

Configuration		**Distance to the Walking Path**		
		1 m	**5 m**	**10 m**
Source Intensity	100 μCi (127 cps)	90%	20%	10%
	300 μCi (380 cps)	100%	90%	20%
	500 μCi (626 cps)	100%	100%	50%
	1,000 μCi (1,220 cps)	100%	100%	90%
	2,000 μCi (2,354 cps)	100%	100%	100%

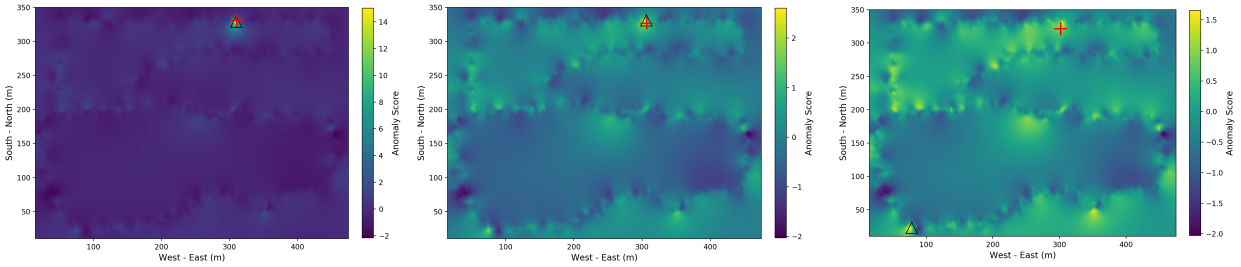
To better illustrate the performance of the proposed method under different source intensities and distances, additional sample cases are shown in Fig. 5.14. For each row, the source intensities are the same, but the distances of radioactive source to the walking path increase. For each column, the distances to the walking path are the same, but the source intensities increase. When source intensity is too low (e.g., 100 μCi) or the source is too far away from the walking path (e.g., 5 meters or 10 meters), the proposed algorithm cannot correctly detect the radioactive sources. But the performance improves when the source intensity is increased. For example, considering the third row of second column and third column of Fig. 5.14, as the source intensity increases to 500 μCi, the proposed algorithm still works even when the radioactive source is placed 10 meter away from the walking paths.

5.4 Summary

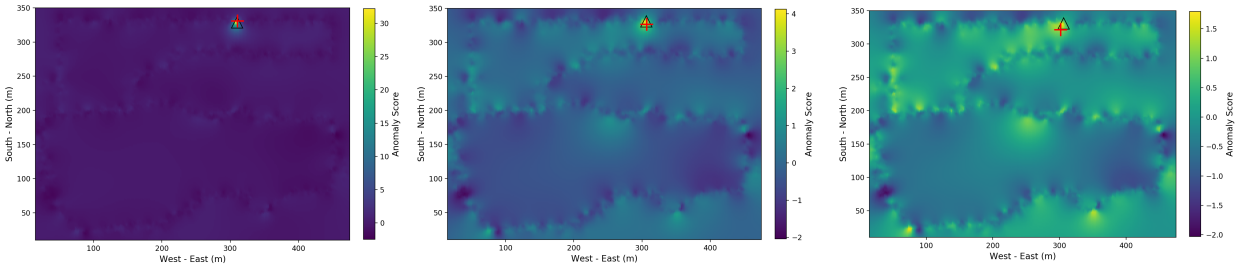
In this chapter, Poisson kriging is implemented to model the geospatial count data measured from mobile sensor networks, which takes the geospatial correlation into account. Based on the estimation,



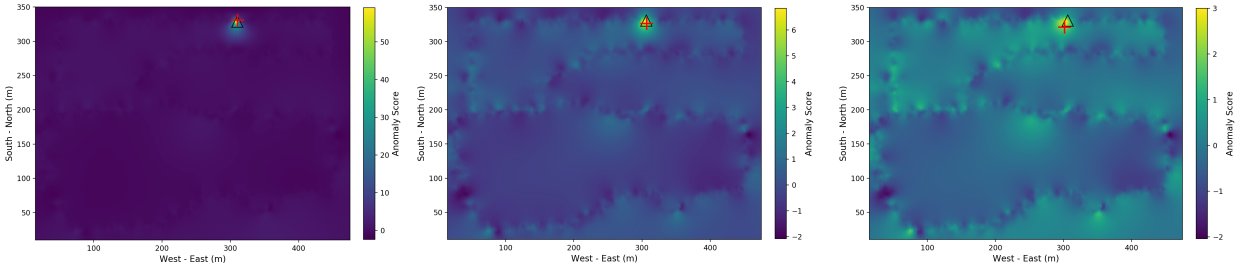
(a) 100μCi source 1 meter away (b) 100μCi source 5 meter away (c) 100μCi source 10 meter away



(d) 300μCi source 1 meter away (e) 300μCi source 5 meter away (f) 300μCi source 10 meter away



(g) 500μCi source 1 meter away (h) 500μCi source 5 meter away (i) 500μCi source 10 meter away



(j) 1000μCi source 1 meter away (k) 1000μCi source 5 meter away (l) 1000μCi source 10 meter away

Figure 5.14: Poisson kriging anomaly detection for selected experiments. Actual radioactive source locations are denoted by red crosses and estimated source locations are denoted by black triangles.

anomalous source detection is performed to detect the potential radioactive sources. The performance of the proposed framework is analyzed using a small mobile sensor network and experimental (measured) data with simulated radioactive sources injected. The results indicate that the proposed algorithm can find the anomalous radioactive source with extremely high accuracy if the source is close enough to the walking path (e.g., 1 meter away from the walking path) or if the radioactive source is strong enough (e.g., 500 μCi or above). For cases where the radioactive source is placed 5 meters away from the walking paths, the number of successful cases increases rapidly when the source intensity increases. The model yields poor performance when the sources studied are more than 10 meters away from the walking paths.

On the other hand, there are several uncertainty factors in this study. The radiation level that is measured using a mobile sensor is significantly affected by weather condition, detector's shielding condition, and moving speed. The correlation between those factors and the background radiation level is not analyzed in this work. Also, the current Poisson kriging model assumes that the background radiation level has the same mean value among the region of interest. A more advanced hierarchical model [84] can also be used to avoid the uniform mean assumptions. In addition, the proposed algorithm is only tested using a small mobile sensor network for weak radioactive sources. In practical applications, the computational complexity should be considered and the optimal threshold for anomalous source detection should be determined using much larger experimental data.

Poisson kriging is useful for estimating nuclear radiation levels given measurements available at only parts of the studied area. It can also be used to identify smaller regions that may contain radioactive sources. But its performance on accurately locating the radioactive sources are limited by many factors. In the next chapter, a source localization technique based on maximum likelihood estimation is proposed, which provides more accurate results for source localization.

Chapter 6

Parameter Estimation for Anomalous Source Localization

6.1 Introduction

Chapter 5 introduced Poisson kriging to estimate background radiation levels and detect anomalous source. Considering the geospatial correlation between radiation measurements, Poisson kriging can correctly estimate background radiation levels using measured radiation data from nearby locations and reach fairly high accuracy in detecting radioactive sources when the source intensities are high or the radioactive sources are close to the detectors. Poisson kriging is more useful in determining the existence of radioactive sources than accurately locating radioactive sources. For source localization, other methods are needed.

Using the results from Poisson anomaly score as described in section 4.2 or Poisson kriging model from Chapter 5, the existence of potential radioactive sources can be determined, and further search can be restricted to smaller regions, instead of the entire studied area. This chapter focuses on radioactive source localization and estimation problems. Without relying on active search, kernel density estimation (KDE) and grid search are implemented respectively with maximum likelihood estimation (MLE) on the streaming data to estimate the source location and intensity. The scenarios for stationary sources are discussed in detail. The performance of the proposed algorithms is compared and demonstrated using a simulation study and a small-scale experiment with simulated radioactive sources injected.

6.2 Background and Theory

6.2.1 Parameter Estimation for Source Localization

As discussed in Chapter 2 and Chapter 5, the measured count rate by a detector is assumed to follow the Poisson distribution. The probability that a detector collects m counts in a unit time with expectation λ is:

$$p(x = m) = \frac{\lambda^m}{m!} \cdot e^{-\lambda}. \tag{6.1}$$

Suppose that some radioactive source with source intensity s_0 is placed at location $\boldsymbol{r}_0$, where $\boldsymbol{r}_0 = (x_0, y_0) \in \mathbb{R}^2$, and the i-th detector is placed at position $\boldsymbol{r}_i = (x_i, y_i)$. The distance of the i-th detector to the radioactive source is $d_i = ||\boldsymbol{r}_i - \boldsymbol{r}_0|| = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$. Then the average count rate at distance d_i should be:

$$\lambda_i = s_i + b_i = \frac{s_0}{d_i^2} \cdot \exp\left(-\int_{\boldsymbol{r}} \mu d\boldsymbol{r}\right) + b_i, \tag{6.2}$$

where μ is the linear attenuation coefficient[1] (m^{-1}), $\exp\left(-\int_{\boldsymbol{r}} \mu d\boldsymbol{r}\right)$ is due to the attenuation effect, and b_i is the average background radiation count rate at position $\boldsymbol{r}_i$, which are influenced by many factors, such as location and weather condition.

In the above equation, there is assumed to be only one stationary radioactive source which is isotropic, and the radiation signals from background and radioactive sources are independent. In addition, all detectors are assumed to have the same properties, such as detector efficiency and detector face area. Since the size of the detector and radioactive source are very small compared with the distance between radioactive source and detector, both detector and source are considered as points. Under the above assumptions, the probability of a detector at position $\boldsymbol{r}_i$ collecting m_i count rate is:

$$p(x_i = m_i; \boldsymbol{r}_i, s_0, \boldsymbol{r}_0, b_i) = \frac{(b_i + s_i)^{m_i}}{m_i!} \cdot e^{-(b_i + s_i)}. \tag{6.3}$$

Considering the case where N mobile detectors are deployed in the studied area, the positions of all detectors at time t are denoted by $\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_N$, and the measured gross count rates for each detector are denoted by $m_1, m_2, \cdots, m_N$. Since in real cases it is almost impossible to know the influence of attenuation for all detectors, the influence of attenuation in Eq. 6.2 is ignored. Suppose that for each potential radioactive source, only a relatively small region (for example, 40 m $\times$ 40 m square region in this work) is considered for source localization. Then it is reasonable to assume that the expected background radiation b_i is nearly constant over the considered area over a short time period, which is denoted by b_0. The likelihood function under the above assumptions is:

$$\begin{aligned} \mathcal{L}(s_0, \boldsymbol{r}_0, b_0) &= \prod_{i=1}^{N} p(x_i = m_i; \boldsymbol{r_i}, s_0, \boldsymbol{r}_0, b_0) \\ &= \prod_{i=1}^{N} \frac{(b_0 + s_i)^{m_i}}{m_i!} \cdot e^{-(b_0 + s_i)}. \end{aligned} \tag{6.4}$$

[1] See Table 2.1 for more details about linear attenuation coefficient.

And the log-likelihood function can be easily derived:

$$\begin{aligned}\ell(s_0, \boldsymbol{r}_0, b_0) &= \sum_{i=1}^{N} [m_i \cdot log(b_0 + s_i) - (b_0 + s_i) - log(m_i!)] \\ &= \sum_{i=1}^{N} \left[m_i \cdot log(b_0 + \frac{s_0}{||\boldsymbol{r}_i - \boldsymbol{r}_0||^2}) - (b_0 + \frac{s_0}{||\boldsymbol{r}_i - \boldsymbol{r}_0||^2}) - log(m_i!) \right].\end{aligned} \tag{6.5}$$

Equation 6.5 is a function of source intensity s_0, source position $\boldsymbol{r}_0$, and average background radiation level b_0. Since the estimation of parameters $(s_0, \boldsymbol{r}_0, b_0)$ cannot be directly computed using standard MLE method due to multiple local maxima, grid search is needed to find the best combination of s_0, $\boldsymbol{r}_0$, and b_0 such that Eq. 6.5 is maximized [68, 90].

Grid search over s_0, $\boldsymbol{r}_0$, and b_0 is computationally expensive, especially when the region to be searched is large or there are many discrete positions to be searched. It can be shown that[2], for a given source position $\boldsymbol{r}_0$, the profile log-likelihood function $\ell_{\boldsymbol{r}_0}(s_0, b_0)$ is concave in s_0 and b_0. Based on this, once the position of the radioactive source can be determined, the source intensity can be easily calculated through standard MLE. This optimization problem can be solved through kernel density estimation with maximum likelihood estimation or grid search with maximum likelihood estimation. The following sections will illustrate this idea more precisely[3].

6.2.2 Kernel Density Estimation with Maximum Likelihood Estimation

As discussed in Section 6.2.1, locating the radioactive source and estimating the source intensity can be separated into two steps. The first step is to estimate the location of radioactive source. In this section, kernel density estimation (KDE) is proposed to estimate the location of radioactive sources. Then, standard MLE is applied to estimate source intensities.

As described in section 2.5, the idea behind KDE is to estimate the probability density based on given data using a kernel $K(\boldsymbol{r}_i, \boldsymbol{r})$, where $\boldsymbol{r}_i$ is the location of the i-th detector and $\boldsymbol{r}$ is the location to be estimated. Here, KDE is implemented in 2D instead of 1D. Based on detector positions $\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_N$ and count rate measurements $m_1, m_2, \cdots, m_N$, for any given position $\boldsymbol{r}$, the estimated probability density is:

[2] See reference [68] or Appendix H for detailed proof.

[3] It might be necessary to point out the difference between this work and the work from Liu et al. [39]. In their implementation, they mainly used method based on MLE to model the background radiation. The spatial and temporal factors were treated as two independent Poisson random variables. There was no explicit consideration of radioactive sources in their model. But in this work, the spatial variation is ignored. Instead, this work treats radiation from background radiation and radioactive sources as two independent variables and tries to locate the radioactive source and estimate source strength through iterative processes, which is quite different from their work.

$$\hat{f}(\boldsymbol{r}) = \frac{1}{C}\sum_{i=1}^{N} m_i K(\boldsymbol{r}_i, \boldsymbol{r}), \tag{6.6}$$

where C is the normalizing constant such that $\sum_{\boldsymbol{r}} \hat{f}(\boldsymbol{r}) = 1$, and $K(\boldsymbol{r}_i, \boldsymbol{r})$ is the kernel function. In this work, a Gaussian kernel $K(\boldsymbol{r}_i, \boldsymbol{r}) = \exp\left(-\frac{||\boldsymbol{r}_i - \boldsymbol{r}||^2}{2\sigma^2}\right)$ with bandwidth σ is used.

The estimated density $\hat{f}(\boldsymbol{r})$ is assumed to represent the likelihood of radioactive source. Then, the position with the highest density, $\hat{\boldsymbol{r}} = \arg\max_{\boldsymbol{r}} \hat{f}(\boldsymbol{r})$, is the most likely source location based on KDE. After determining the most likely source location, source intensity can be calculated using standard MLE method, which is essentially maximizing the profile log-likelihood function $\ell_{\boldsymbol{r}_0}(s_0, b_0)$ given the source location $\boldsymbol{r}_0$. This process is illustrated in Fig. 6.1. Section 2.7 introduced gradient ascent and Newton-Raphson methods for optimization, and Appendix I contains the details for these two methods.

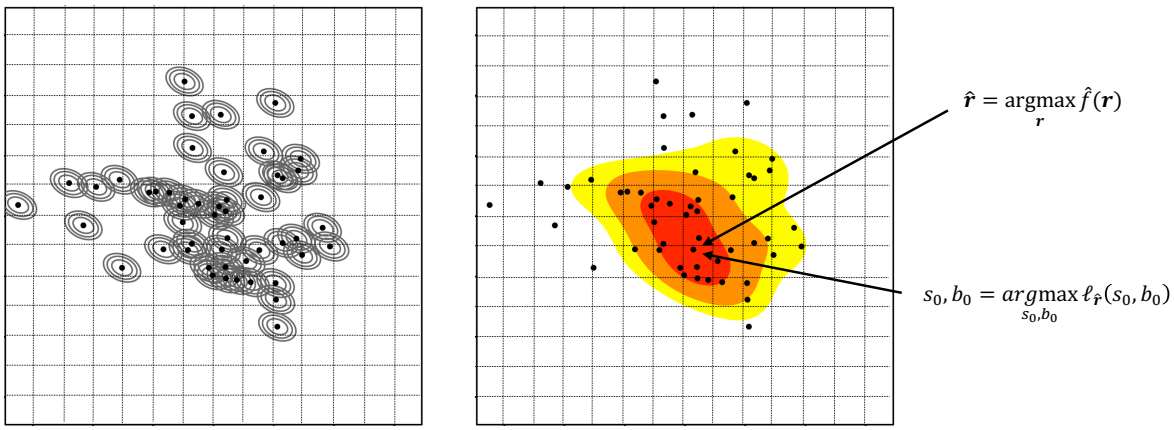


Figure 6.1: Illustration of KDE with MLE for source localization and strength estimation.

6.2.3 Grid Search with Maximum Likelihood Estimation

Section 6.2.2 proposed the application of KDE with MLE for source localization and strength estimation. In addition to estimating source location using KDE, a more straightforward solution is to calculate the source intensity s_0 and background radiation b_0 that maximize the profile log-likelihood function $\ell_{\boldsymbol{r}_0}(s_0, b_0)$ at all possible source locations[4] $\boldsymbol{r}_0$. Then the location that has the highest likelihood value will be the most likely source location. This leads to the grid search with MLE method.

The process of grid search with MLE is illustrated in Fig. 6.2. After the suspicious region is identified, the entire region will be divided into small grids. Then, each node $\boldsymbol{r}_0$ is assumed to be the location of

[4]In cases when the background radiation can be estimated in advance, only the source intensity s_0 that maximizes the likelihood function is needed, which can simplify the calculation.

radioactive sources s_0. The profile log-likelihood function $\ell_{\boldsymbol{r}_0}(s_0, b_0)$ is calculated and source intensity s_0 and background radiation level b_0 that maximize $\ell_{\boldsymbol{r}_0}(s_0, b_0)$ is determined. This process is repeated for every node[5]. Finally, the location with the highest likelihood, $\hat{\boldsymbol{r}_0} = \arg\max_{\boldsymbol{r}_0} \ell(s_0, \boldsymbol{r}_0, b_0)$, is the most likely source location based on grid search.

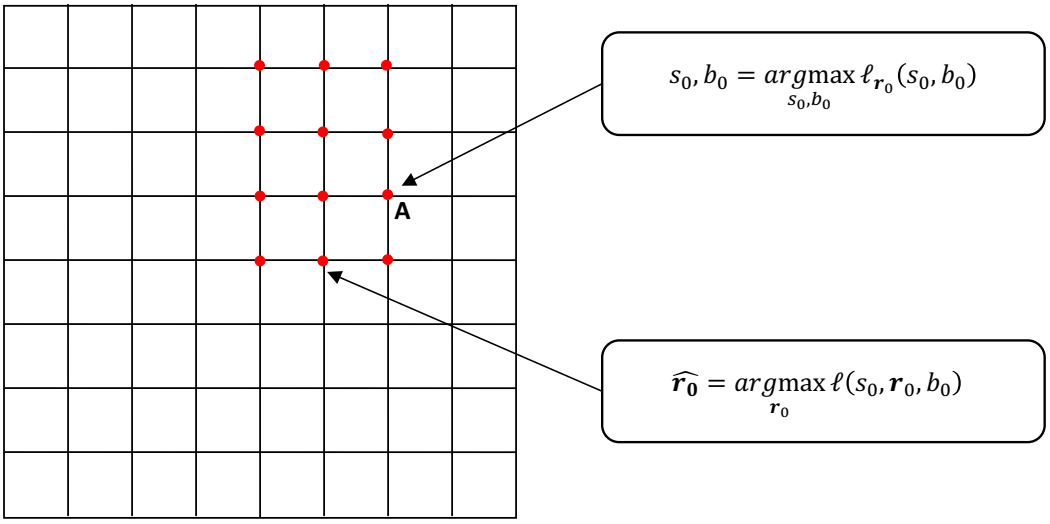


Figure 6.2: Illustration of grid search with MLE for source localization and strength estimation.

6.2.4 Implementation

The implementations of KDE with MLE and grid search with MLE for source localization and estimation are quite similar with each other. A flowchart of this process is shown in Fig. 6.3.

In practical cases, the calculations are assumed to begin at time t when there are signs that some potential radioactive sources exist and the suspicious region that may contain radioactive sources is identified. For example, k-sigma method [6] or Poisson kriging model can be applied first to determine the existence of radioactive sources. After this step, the system will continue collecting data in the identified region and KDE with MLE or grid search with MLE can be applied to locate the radioactive sources. Both procedures can be applied continuously as new measurements within the identified region are collected by detectors. The is referred as the calculation steps. With limited number of mobile detectors, especially when number of detectors is small or the entire region to be monitored is large, for each time t, there might be no measurements collected within the identified suspicious region. In this case, there is no need to repeat the calculation since there is no information updated. In this way, after one calculation step, several seconds might pass in actual time before the next calculation step is carried out.

[5]The size of the grids will influence the source localization precision. When the gird size is too large, the algorithm may miss the exact location and give large source location error. When the grid size is too small, the precision can be high, but it will take more computation since the same process needs to be repeated on each node.

Theoretically, the calculation can be stopped when there is no significant change in the estimated source positions. As will be seen later, this usually takes a large number of calculation steps to converge, especially when the number of detectors is small. In real cases, as the calculation continues and the potential sources are approximately identified, security officers who carry more sensitive detectors can be sent to the suspicious region to detect the radioactive sources.

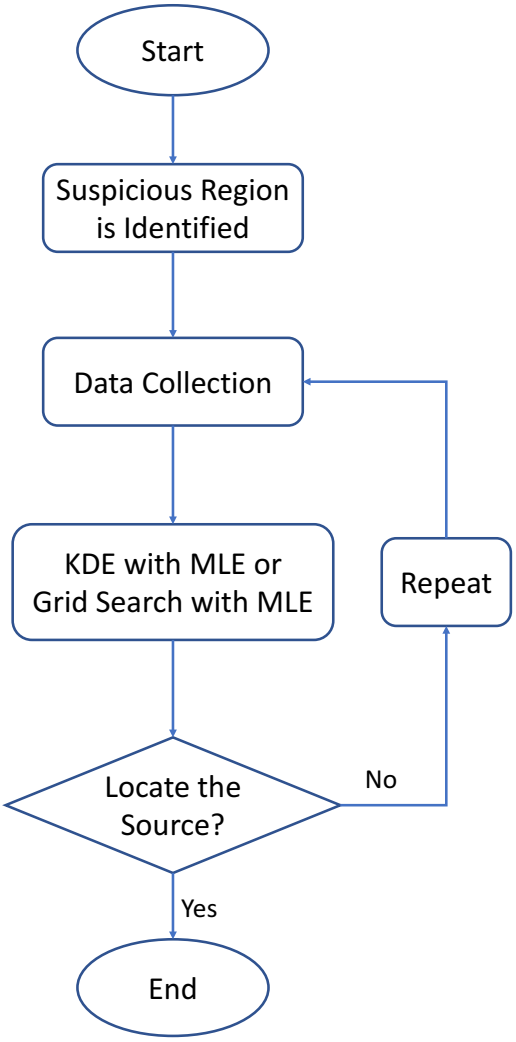


Figure 6.3: Flowchart of KDE with MLE and Grid Search with MLE.

6.3 Results and Discussion

In this work, only stationary radioactive sources are considered, in which the radioactive source is assumed to stay at some position for a relatively long period of time such that the mobile detectors can continuously

collect data from the region surrounding the radioactive source. To test the proposed source localization and estimation methods, the simulated data and experimental data as described in Chapter 3 are used. The performance of KDE with MLE and grid search with MLE is presented and compared.

6.3.1 Simulation Results

In the simulation, the scenario where 50 mobile D3S detectors move in the city is simulated, which is depicted in Fig. 6.4. The simulated map and parameter settings are the same as described in section 3.3. In the simulation, multiple stationary sources are randomly distributed with or without shielding. Four typical source locations are denoted by red circles with numbers 1, 2, 3, and 4.

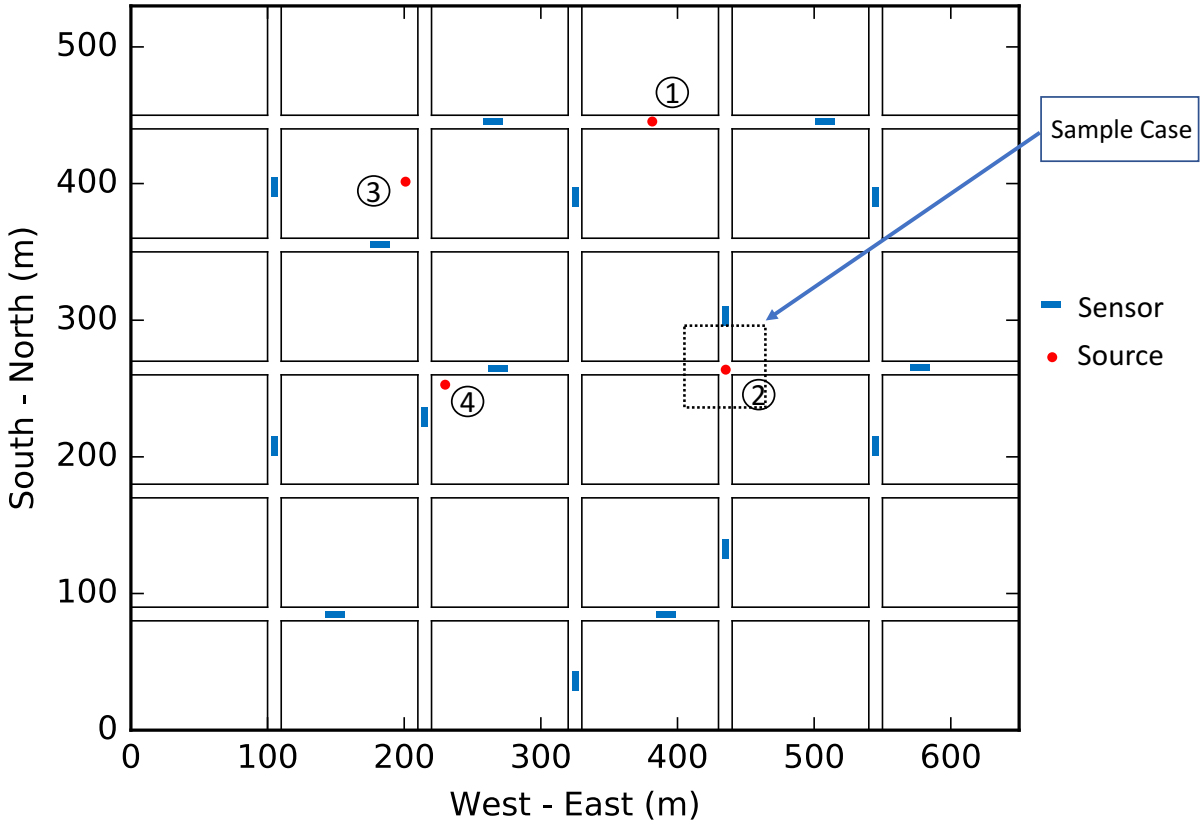


Figure 6.4: Illustration of simulated city map.

Results: KDE with MLE Approach

The goal of KDE is to estimate the probability density in the selected region to locate the radioactive source. As time goes on, more measurements will be collected in the study area. The peak region of KDE

tends to be closer to the actual source location.

To better illustrate the process, a sample case where the radioactive source is located at position 2 in Fig. 6.4 is chosen. Assuming that after anomaly detection, the region surrounding position 2 is suspected to contain radioactive source, and a square region (40 m $\times$ 40 m) is selected for further analysis. Using the procedures described in section 6.2.2 and Fig. 6.3, the results of KDE with MLE are shown in Fig. 6.5.

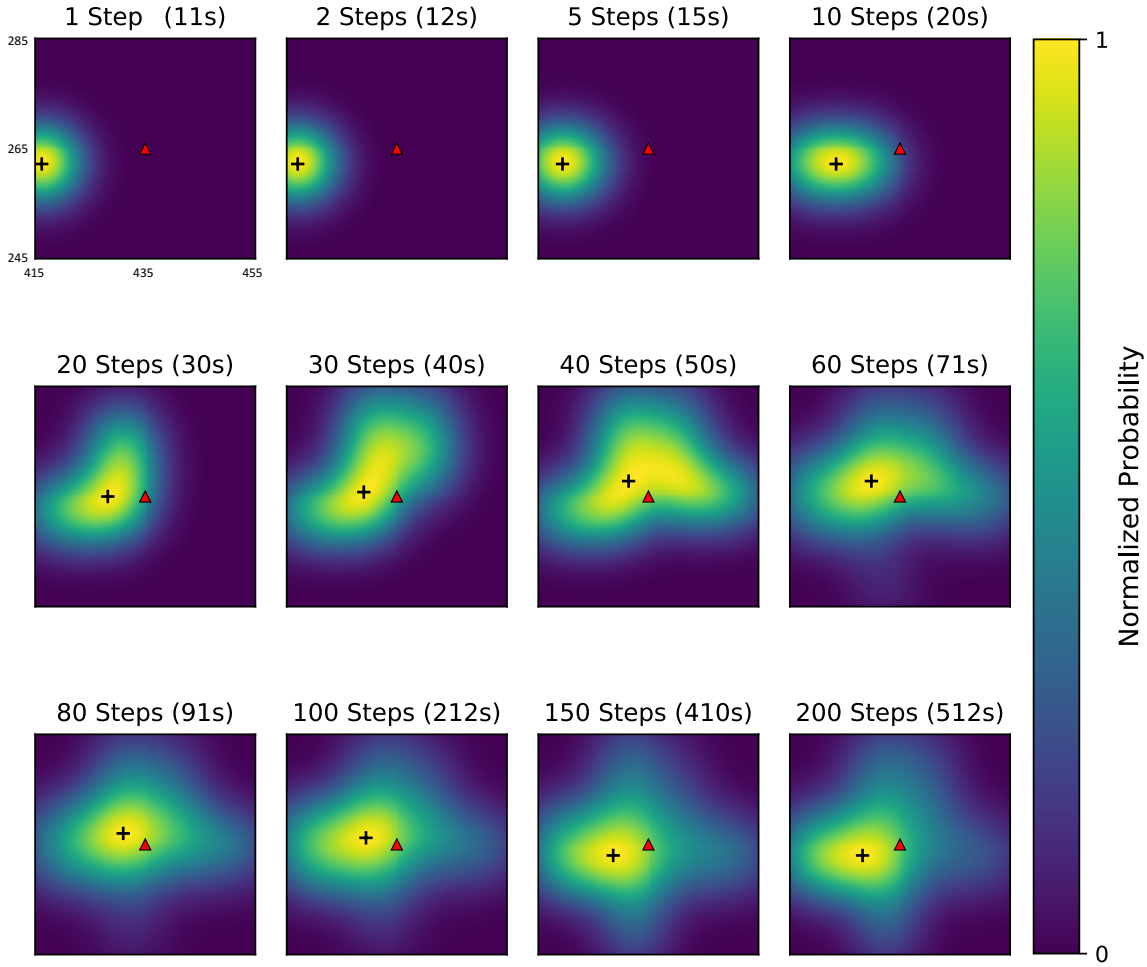


Figure 6.5: Results of KDE approach with bandwidth 5 m when the source is placed at the intersection of the roads (scenario with source number 2 in Fig. 6.4). The actual source location is denoted by the red triangle while the peak location of probability density is denoted by the dark cross.

Figure 6.5 shows the results for the first 200 calculation steps for the sample case using Gaussian kernel with bandwidth of 5 meters. Calculation steps are given on the top of each figure, while the actual simulation time in seconds is included in the parentheses. Since in some time steps, no detector enters the selected suspicious region and hence there is no change in the configuration to be studied. Hence, calculations are not carried out for those time steps. Therefore, the actual simulation time is much larger than calculation

steps. The location of the injected radioactive source is denoted by the red triangle, and estimated source location is denoted by the dark cross, which has the largest probability density calculated using the KDE approach. The background color represents the calculated probability density from the KDE approach. It is normalized to have the largest value of 1 and smallest value of 0. Clearly, as time increases, with more data being collected, the estimated source location gets closer to the actual source location. But KDE with MLE approach cannot find the exact location of the radioactive source even after 200 calculation steps.

To quantify the overall performance of KDE with MLE, Fig. 6.6 shows the relation of the estimated source location error versus calculation steps. The source location error is calculated using the Euclidean distance between the estimated source location and the actual source position. The errors are calculated for all simulated cases corresponding to to four different source positions depicted in Fig. 6.4[6]. The average source localization error and corresponding standard error of four different Gaussian kernel bandwidths (5 m, 10 m, 15 m, and 20 m) are shown in the figure.

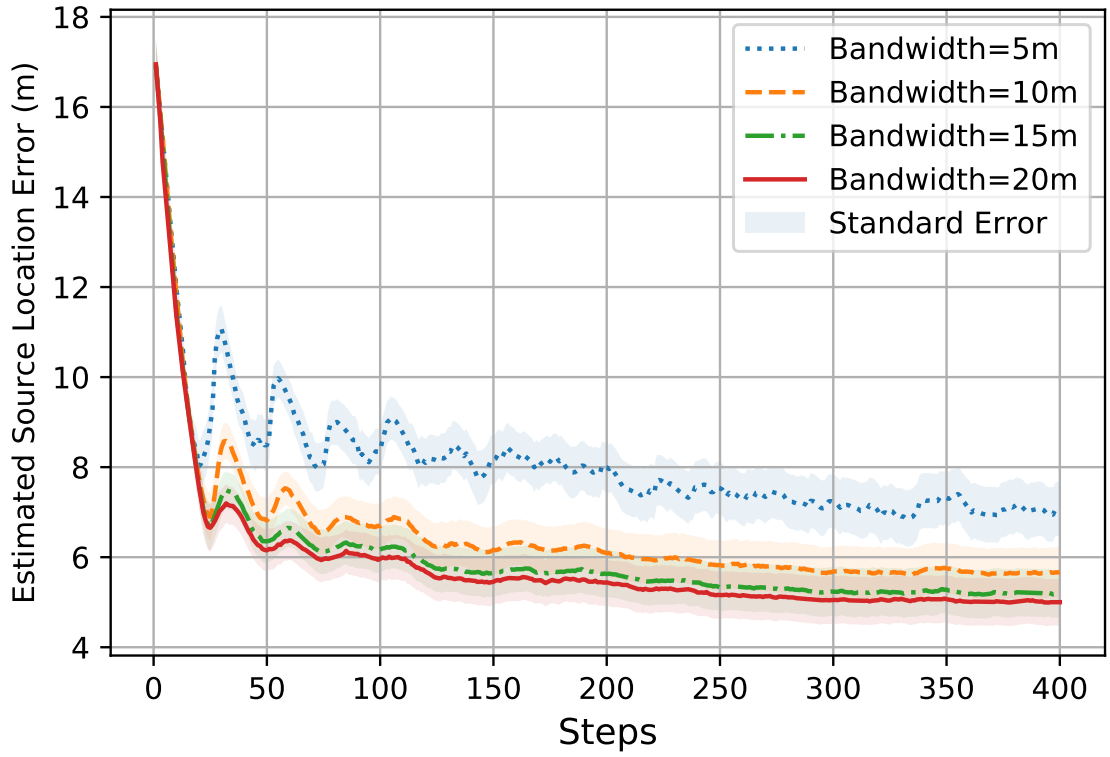


Figure 6.6: Change of estimated source location error with different Gaussian kernel bandwidths (σ in $K(\boldsymbol{r}_i, \boldsymbol{r}) = \exp\left(-\frac{||\boldsymbol{r}_i - \boldsymbol{r}||^2}{2\sigma^2}\right)$). Results are averaged for 4 typical scenarios depicted in Fig. 6.4.

In Fig. 6.6, in all cases, the source localization error gradually decreases as the calculation steps increase

[6]Figure 6.4 only includes four typical source locations. In the actual simulation, there are more than 10 different source locations simulated for each typical location from 1 through 4.

with some fluctuation at the start of the process. Increasing the bandwidth from 5 meters to 10 meters can dramatically reduce the source localization error. But as the bandwidth keeps increasing, after 15 meters, there is little improvement. After 400 calculation steps, KDE with bandwidth of 20 meters can reduce the source localization error to be around 5 meters.

Results: Grid Search with MLE Approach

With KDE approach, although the source localization error can be reduced to be around 5 meters, this takes around 400 calculation steps, which takes about $1,000$ seconds in simulation time under current configuration (36 city blocks with 50 mobile detectors). To further reduce the source localization errors with fewer calculation steps, grid search with MLE is applied. Similar to Fig. 6.5, the results of a sample case for scenario corresponding to source number 2 in Fig. 6.4 for the first 90 calculation steps are shown in Fig. 6.7. In the figure, background color corresponds to the calculated likelihood value from grid search with MLE approach. The color is normalized to be within the range of [0, 1]. The location with the highest likelihood value is denoted by dark cross, which is the estimated source location from the grid search method.

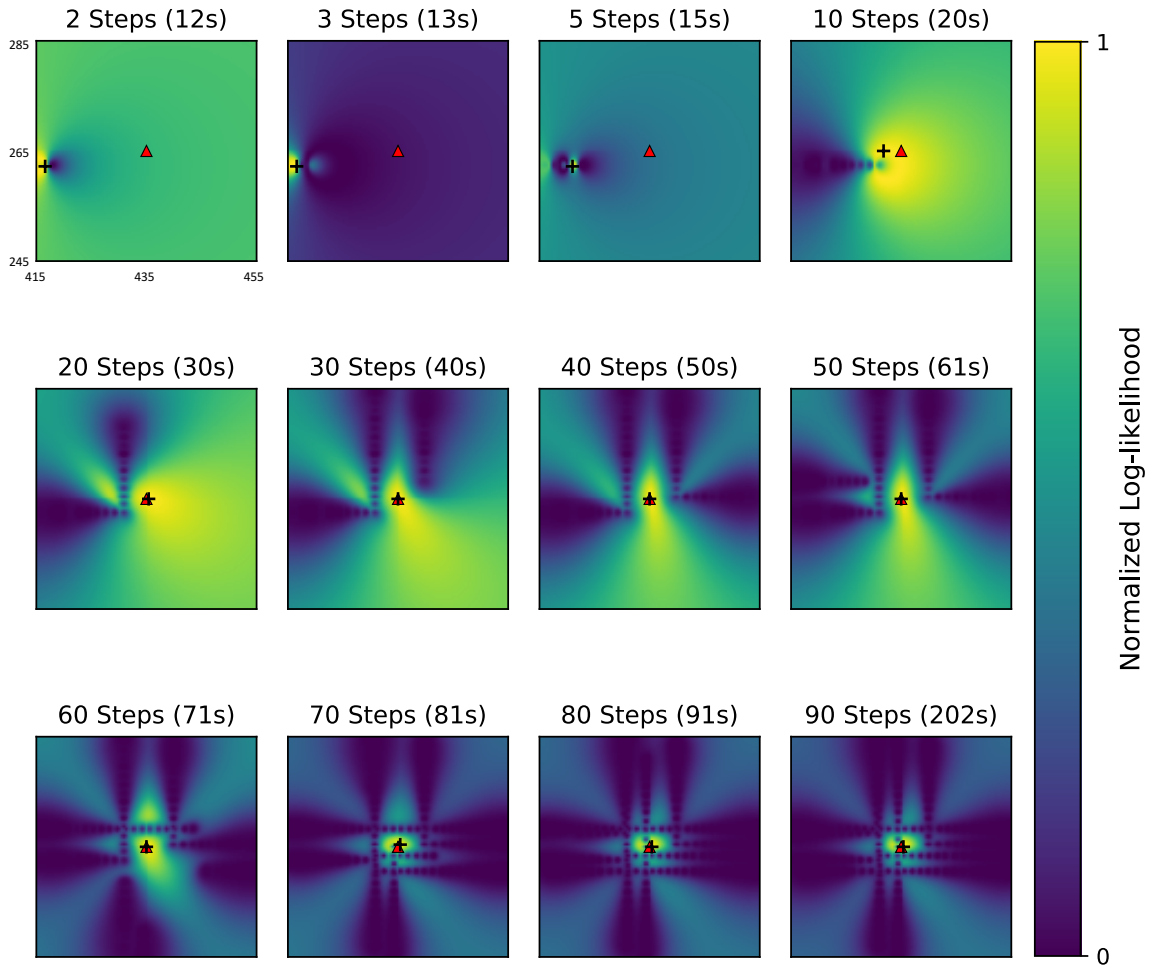


Figure 6.7: Results of the grid search with MLE when source is placed at the intersection of roads (scenario with source number 2 in Fig. 6.4). The background color corresponds to the normalized likelihood value (scaled to be between 0 and 1). The actual source location is denoted by red triangle while the peak location of normalized likelihood value is denoted by dark cross.

It is obvious that for the same sample case, grid search with MLE approach performs much better compared with the KDE approach. The estimated source location is almost identical (within 1 m source localization error) to the actual source location after 20 calculation steps (30 seconds in simulation time). To show the overall performance under different cases, the averaged source localization error of grid search for four sources is shown in Fig. 6.8. For comparison, Fig. 6.8 also includes the results of the KDE approach with bandwidth equals 10 meters using dotted line.

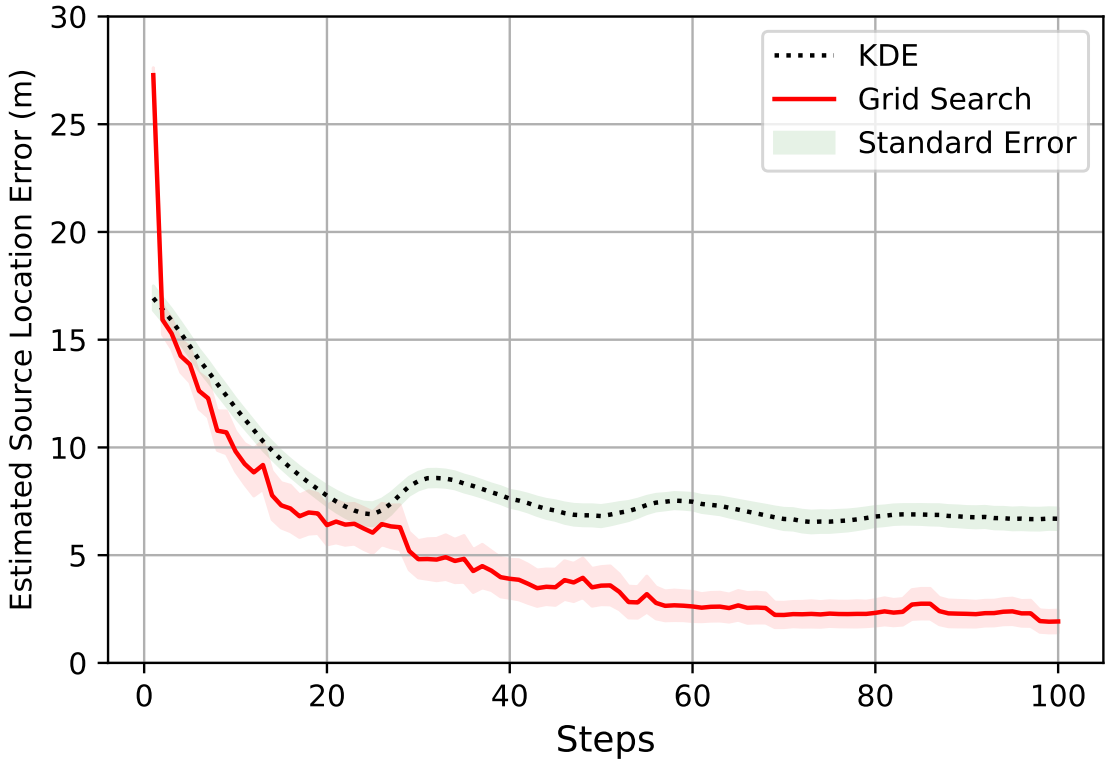


Figure 6.8: Change of estimated source location error as a function of steps for KDE with MLE approach and grid search with MLE approach. Results are averaged for 4 source scenarios depicted in Fig. 6.4.

As shown clearly in Fig. 6.8, in the simulation study, the grid search method outperforms the KDE method by reducing the source localization error as well as decreasing the calculation steps. Grid search with MLE approach can reduce the source location error to be within 5 meters in about 30 calculation steps, while KDE will take more than 400 calculation steps. After 100 calculation steps, grid search with MLE approach can reduce the source localization error to be around 2 meters, while KDE can only reduce it to be around 7 meters.

In addition to source localization error, another metric is the source intensity estimation error. Figure 6.9 plots the percentage of estimated source intensity error as the calculation step increases. After 100 calculation steps, the grid search with MLE approach can reduce the source intensity estimation error to be within 10%, while KDE with MLE approach has much larger estimation error (around 60%).

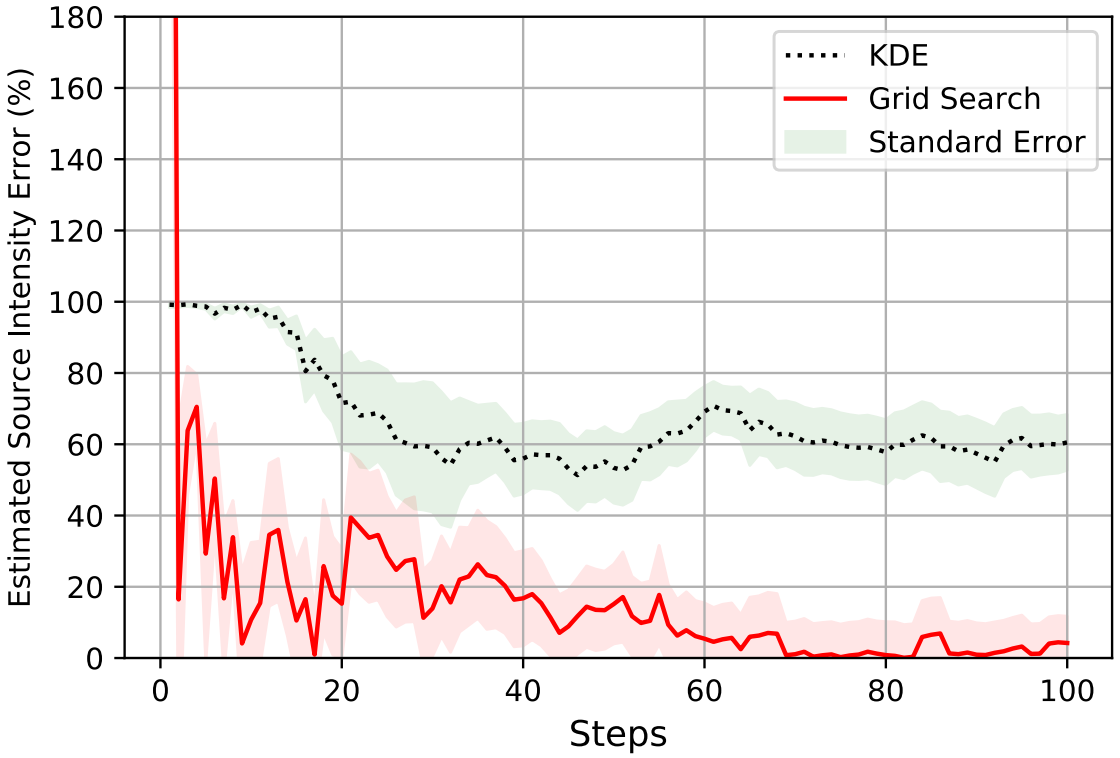


Figure 6.9: Change of source strength estimation error with KDE approach and grid search approach. Results are averaged for 4 typical scenarios depicted in Fig. 6.4.

Although grid search with MLE approach works well for estimating the source intensity, it is worth mentioning that this is due to the fact that in the simulation, the attenuation effect is quite low and the simplified model can still model the radiation transport process without a large error. In real cases, however, considering the existence of buildings, the simplification made for the radiation transport model in section 6.2.1 will dramatically influence the proposed method's performance. To reduce the estimation error, more accurate physical models, especially the model considering the shielding effect around the source, need to be applied.

6.3.2 Results for Sources Injected in Measured Background Radiation Data

Simulation results for KDE and grid search approaches were presented in the previous sections. To better quantify the performance, the proposed methods are then applied to the experimental (measured) data with injected (simulated) sources as described in Section 3.4. During the experiment, the simulated radioactive sources with source intensity ranging from 100 μCi to $2,000$ μCi are placed at 10 different positions (shown in Fig. 3.6). The distance from the sources to walking paths is varied from 1 meter to 10 meters. Following

the steps described in the previous sections, the average source localization errors[7] for KDE approach and grid search approach after 80 calculation steps are shown in Table 6.1 and Table 6.2.

Table 6.1: Average Source (^{137}Cs) Localization Error for KDE Approach

Configuration		**Distance to the Walking Path**		
		1 m	**5 m**	**10 m**
Source Intensity	100 μCi (127 cps)	9.72 m	10.73 m	13.77 m
	300 μCi (380 cps)	9.63 m	10.68 m	13.77 m
	500 μCi (626 cps)	9.50 m	10.68 m	13.77 m
	1,000 μCi (1,220 cps)	9.41 m	10.63 m	13.77 m
	2,000 μCi (2,354 cps)	9.37 m	10.59 m	13.68 m

Table 6.2: Average Source (^{137}Cs) Localization Error for Grid Search Approach

Configuration		**Distance to the Walking Path**		
		1 m	**5 m**	**10 m**
Source Intensity	100 μCi (127 cps)	2.27 m	7.53 m	17.31 m
	300 μCi (380 cps)	0.76 m	4.03 m	16.76 m
	500 μCi (626 cps)	0.62 m	2.22 m	10.10 m
	1,000 μCi (1,220 cps)	0.50 m	1.76 m	3.99 m
	2,000 μCi (2,354 cps)	0.52 m	0.62 m	2.05 m

Compared with the simulation results shown in Fig. 6.8 (for the simulated background radiation case), KDE approach performs much worse for the measured background radiation data. The main reason is that, during the experiment, the distance of mobile detectors to the injected radioactive source is limited to be more than 1 meter, 5 meters, or even 10 meters, and the spatial distribution of those measurements are not symmetric for the injected source. These factors will bias the peak of KDE toward one single direction, instead of centering around the injected source, which leads to the worse performance of KDE compared with the simulated data case.

Grid search approach, on the other hand, performs consistently well for the experiment. It can reduce the source localization error to be within 2.5 meters when 100 μCi radioactive source is injected 1 meter from

[7]There are certain cases, for example, when the 100 μCi source is placed 10 meters away from the walking path, the source localization error is very high and it is actually due to the fact that the nearby detectors cannot measure any useful signal. This is because that the source intensity is too low and the distance is too large. These cases can be totally ignored when evaluating the performance of the proposed algorithm for source localization.

the walking path within 80 calculation steps. When the source is injected 5 meters away, it can reduce the source localization error from 7.53 m to 2.22 m when source intensity is increased from 100 μCi to 500 μCi. For the most difficult case where radioactive source is injected 10 meters away, grid search with MLE can still correctly identify the radioactive sources with gradually decreasing source localization errors (decreases from 17.3 m to 2.05 m). This shows the robust performance of grid search with MLE approach.

6.4 Summary

The results of anomaly detection or background radiation estimation are used to identify the potential region that might contain radioactive sources. To pinpoint the radioactive sources, in this chapter, KDE and grid search approaches are implemented with MLE to locate the radioactive sources. Results from simulation study and experiment with simulated sources injected in measured background data quantify the performance of the proposed methods. For stationary sources, grid search with MLE can reduce the source localization error to around 2 meters for 500 μCi source placed 5 meters away from walking path.

Although the KDE approach is worse than the grid search approach, it is much easier to implement and faster to calculate compared with the grid search approach, which make it a good choice to estimate prior information about the potential locations of the radioactive source. Grid search can then be performed based on the results from the KDE approach, which can reduce the space to explore and thus reduce the computational need. In addition, multi-resolution grid computation [91] and adaptive grid search [8] have been proposed to reduce the calculation during grid search, which can also speed up the calculation process.

Additionally, the approaches proposed in this chapter are not restricted to cases where radioactive sources are stationary, Appendix J includes some preliminary analysis of the proposed techniques for mobile radioactive sources. Also, the grid search with MLE approach is closely related to Bayesian estimation, Appendix K shows the same approach but from Bayesian perspective, in which different prior information can be incorporated.

Chapter 7

Spectrum-Based Anomaly Detection

7.1 Introduction

Previous chapters systematically studied the application of mobile sensor network for nuclear radiation detection. Simulations and experiments, mobile sensor network systematic assessment, Poisson kriging, and source localization techniques are discussed in detail. In those analyses, only radiation count rate information is used. Radiation spectra, which contain additional information about nuclear radiation, can also be used for nuclear radiation detection. In this chapter, the spectral information is utilized for anomaly detection.

As shown in Fig. 2.3, radiation spectra are simply the histograms of radiation counts binned by channels or energies. All information from background radiation and radioactive sources are contained in the spectra. Radiation detectors have a certain number of channels. For example, the D3S detector has 512 channels while the Ortec 905 NaI detector has 1,024 channels by default. Spectra of randomly sampled 500 spectra simulated using GADRAS for Ortec 905 detector are shown in Fig. 7.1.

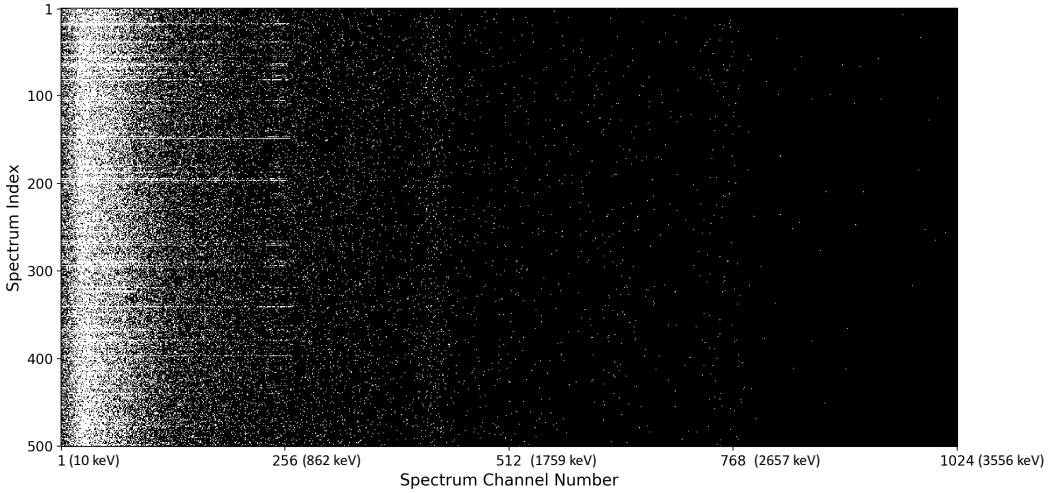


Figure 7.1: Sample spectra with 1024 channels for Ortec 905 NaI detector. The corresponding energies are shown in parentheses. White dots indicate non-zero counts at corresponding channels and black dots indicate zero count.

The spectra in Fig. 7.1 are sampled from the data simulated using GADRAS. The entire simulated dataset contains spectra from background radiation and isotopes[1] such as ^{137}Cs (500 μCi), ^{60}Co (600 μCi), and ^{54}Mn (750 μCi). In the figure, each row represents one spectrum with 1,024 channels, and there are 500 spectra in total. White dots indicate non-zero counts at corresponding channels and black dots indicate zero count at corresponding channels for a given spectrum. It is clear that most records exist in the first several hundred channels and there are few records when the channel number is too large (for example, when channel number is larger than 512). This means that for the most commonly used isotopes, high channel numbers contain little information, and thus they are less useful for anomaly detection. Too many features also increase the problem complexity due to *the curse of dimensionality* [15]. In this chapter, only the information from the first 512 channels for Ortec 905 detectors are used.

As mentioned in section 1.3.1, nuclear radiation detection is essentially an anomaly detection problem. The goal of anomaly detection is to distinguish the anomalous measurements from the normal background measurements. Chapter 3 and Chapter 5 presented several approaches based on radiation count rate data. Using count rate data, since it is only a scalar value, anomaly detection is mostly based on comparing the measured data with some baseline or estimation. However, using spectral information, there are many more possible approaches to address this problem.

This chapter presents the exploratory study of spectrum-based anomaly detection for nuclear radiation data. More specifically, this chapter focuses on two types of spectrum-based anomaly detection problems: unsupervised and semi-supervised anomaly detection[2]. Dimensionality reduction, regression models, and proximity-based models are discussed. The performance of these methods is evaluated using simulated radiation data.

7.2 Dimensionality Reduction and Reconstruction

As discussed in Section 7.1, even after removing the information from less useful channels, there are still 512 channels left. For the purpose of radiation detection using radiation spectra, information from each channel is considered to be a dimension in the problem. The large amount of dimensionality makes it hard to analyze and visualize the data. However, not all dimensions are equally useful for anomaly detection. It is expected that some dimensions help little when performing anomaly detection. This section focuses on the application of dimensionality reduction for anomaly detection. Two typical dimensionality reduction

[1]For the spectra with isotopes, the whole dataset includes the simulation at different distances, ranging from 0.5 m to 20 m.

[2]For unsupervised anomaly detection, there are no labels for the data. The model should be able to distinguish abnormal measurements based on unlabeled data. For semi-supervised anomaly detection, a small part of the data have labels while others are not labeled, or only normal data are labeled (as the focus of this chapter). For more detailed explanation and description, see Reference [16].

techniques, principal component analysis (PCA) and autoencoder, are explored.

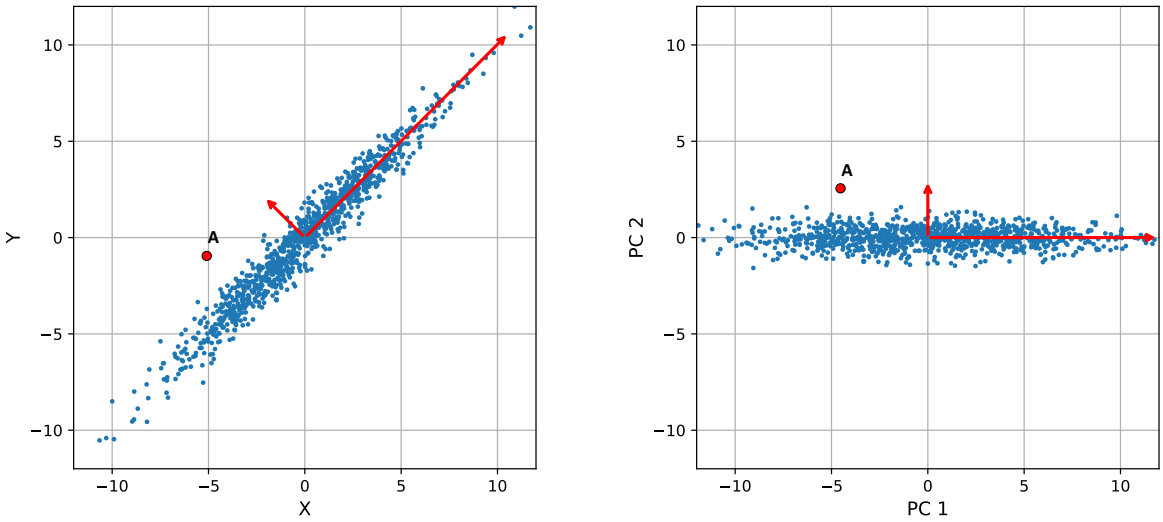
7.2.1 Principal Component Analysis

Principal component analysis (PCA) is a widely used method for feature generation and dimensionality reduction. Suppose that there are n spectra, and each spectrum has m dimensions. The i-th spectrum is denoted by vector $\boldsymbol{x}_i^T$ ($\boldsymbol{x}_i$ is a column vector with 512 dimensions in this work). These n spectra can be represented by a data matrix X with n rows and m columns:

$$X = \begin{pmatrix} \cdots & \boldsymbol{x}_1^T & \cdots \\ \cdots & \boldsymbol{x}_2^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \boldsymbol{x}_n^T & \cdots \end{pmatrix}. \tag{7.1}$$

X contains large amount of spectral information from normal background radiation. The first step of the PCA is to compute the mean values $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_m)$ for each dimension, which correspond to the average background radiation at each channel. Then, for the given data matrix X, by subtracting the mean values in each dimension, all dimensions are made to have a zero-mean, which leads to the centered data matrix X^c. Next, by calculating the covariance matrix of centered data matrix X^c and singular value decomposition (SVD), a series of orthonormal eigenvectors, or principal components, $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$, can be extracted from the input data X^c. All principal components are sorted according to the corresponding variance (eigenvalue) explained by each principal component, which is denoted by $\lambda_1, \lambda_2, \cdots$. Hence, the first principal component explains the largest variance. The second principal component is orthogonal to the first one and explains the second largest variance, and so on.

The above process is shown in Fig. 7.2 using a 2D sample dataset. The original data is plotted over X-axis and Y-axis. Abnormal measurement A is within the normal range for both X-axis and Y-axis, which makes it hard to classify point A as outlier only based on the measured values along these two axes. On the other hand, the original data can be better represented over new axes denoted by red arrows, which correspond to the two principal component directions calculated using the PCA. If the original data are projected onto the two principal component directions, variance on the first principal component direction is much larger than on the second principal component direction. For dimensionality reduction purpose, the original data can be represented using only the projection value on the first principal component direction with minimal information loss.



(a) 2D sample data.

(b) Projection on first two principal component directions.

Figure 7.2: Illustration of PCA on 2D sample data. Sample abnormal measurement A is denoted by red dot.

To use the PCA for anomaly detection, one choice is to use the first k principal components to reconstruct the data [8,32]. Given a new measured spectrum, the reconstructed spectrum is assumed to be the estimation of background radiation for the given spectrum since the PCA model is constructed using background radiation spectra. The difference between the original data and the reconstructed data is then used as the anomaly score. The hyperparameter k plays an important role and the optimal k varies with specific problems. For high-quality spectra, it was suggested that $k = 5$ is a good empirical choice [8]. In that case, the first principal component explains around 55% of total variance, and the top 5 principal components explain around 70% of total variance. However, for low-quality spectra as used in this chapter, it is hard to choose such a small value of k. Figure 7.3 shows the change of cumulative variance explained by various number of principal components for the simulated background radiation spectra. Each principal component only contains little information and the difference among the first several principal components are small, which is quite different from the well-collected spectra [8]. In Fig. 7.3, the first 64 principal components explain 64% of the total variance and the first 256 principal components explain 95% of the total variance.

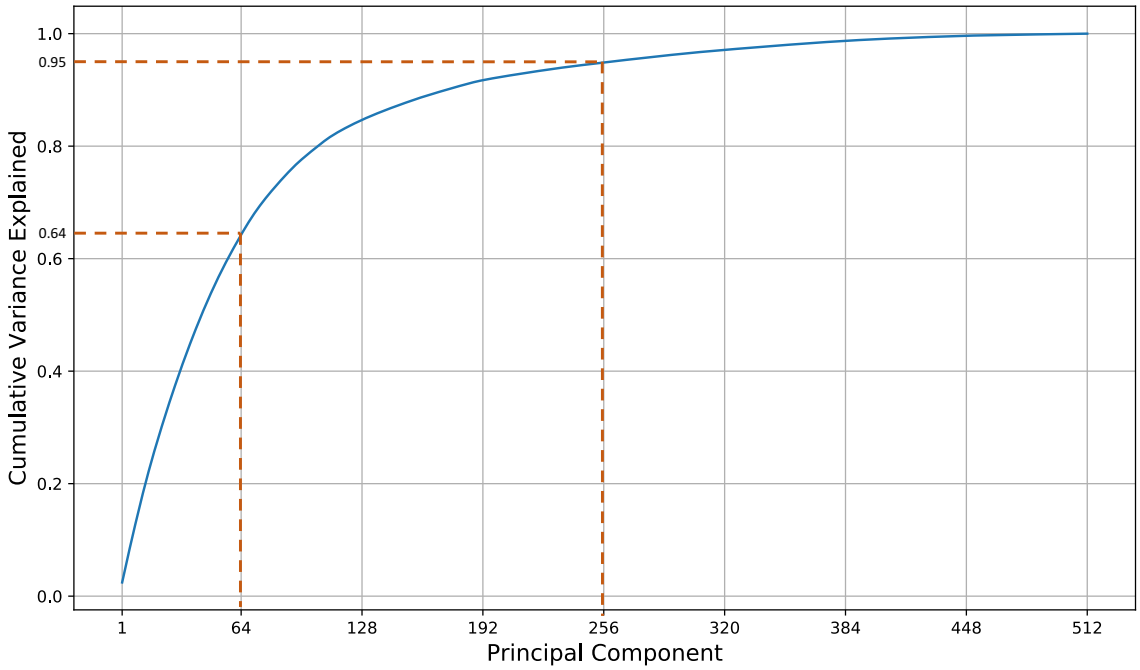


Figure 7.3: Cumulative variance explained by various number of principal components.

In addition to choosing the first k principal components to reconstruct the background spectrum, another choice is by defining a normalized anomaly score consisting of the contribution from all principal components. Mathematically, for a given new measurement $\boldsymbol{x}$, the anomaly score used in this work is defined as [16]:

$$AS_{PCA}(\boldsymbol{x}) = \sum_{i=1}^{m} \frac{|(\boldsymbol{x} - \boldsymbol{\mu}) \cdot \boldsymbol{v}_i|}{\lambda_i}, \tag{7.2}$$

where $\boldsymbol{x}$ represents the new spectrum, λ_i is the eigenvalue, and the corresponding i-th eigenvector is $\boldsymbol{v}_i$.

In this work, the PCA model is constructed only based on the background information. The intuition behind the PCA-based anomaly detection is that, when the new spectrum $\boldsymbol{x}$ which contains radioactive source information is projected onto the principal components space, it tends to have larger deviation on the eigenvectors that have low variance and smaller deviation on the eigenvectors with high variance. While for normal background radiation spectrum, it should have smaller deviation on the eigenvectors that have low variance and larger deviation on the eigenvectors with high variance. Based on this difference, these abnormal spectra can be distinguished from normal background radiation spectra using anomaly score defined in Eq. 7.2. For example, as shown in Fig. 7.2, compared with normal measurements, measurement A deviates much larger in the second principal component direction (PC2) than in the first principal component direction (PC1) compared with normal measurements. Through PCA, measurement A is more easily identified in the

new space than in the original feature space if only the measured values in X-axis and Y-axis are used.

7.2.2 Autoencoder

PCA is used to perform dimensionality reduction and anomaly detection. As a linear transformation[3], PCA is easy to implement and understand. However, linear transformation does not work well for more complex problems. Under such conditions, another technique that is similar to PCA, called autoencoder [70], can be used to perform non-linear dimensionality reduction and reconstruction.

Autoencoders treat the input $\boldsymbol{x}$, for example, radiation spectrum in this work, as the target. The goal is to find a reconstruction function $r(\boldsymbol{x}) = g(h(\boldsymbol{x}))$ through encoder $h(\cdot)$ and decoder $g(\cdot)$ [93]. An autoencoder system is usually a multi-layer deep neural network, which can be trained through back-propagation [70]. Figure 7.4 shows the bottleneck structure of a 5 layer autoencoder system. The input data are represented by input layer. Through the encoder layers, the input data are represented in a lower dimension. The decoder layers are then used to reconstruct the input data. The output layer has the same dimension as the input layer.

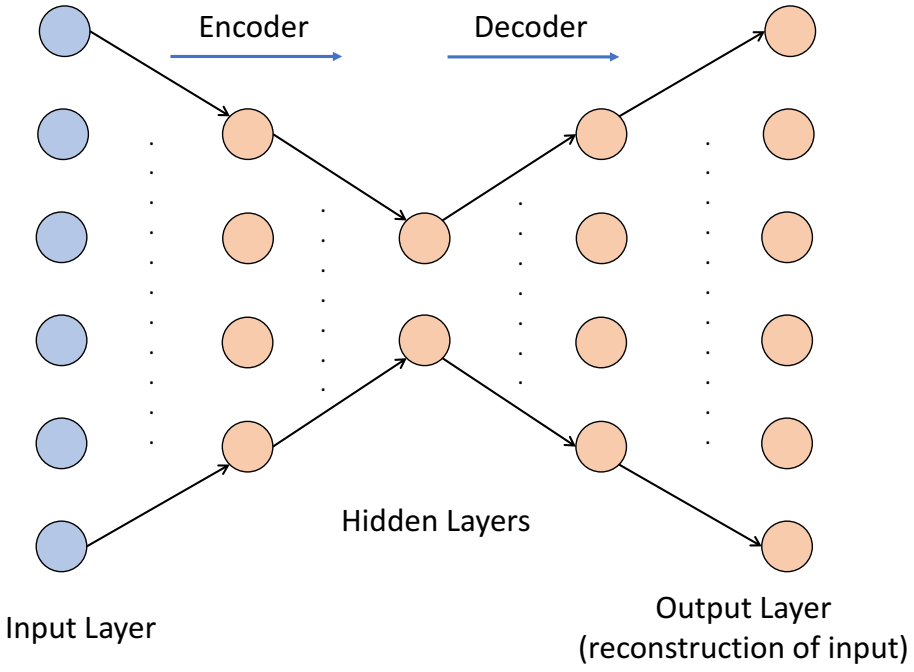


Figure 7.4: Structure of a 5-layer Autoencoder system.

For anomaly detection, the autoencoder is trained on the normal background radiation data. After

[3]There also exists non-linear version of PCA, called kernel PCA. See Ref [92] for more details.

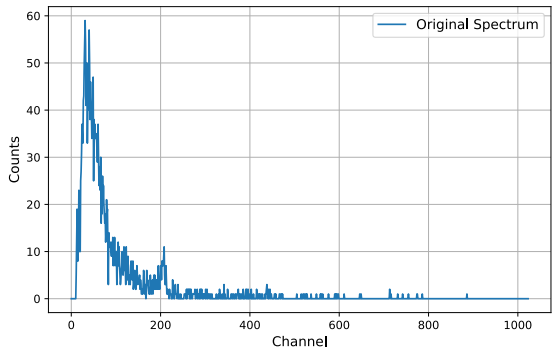
training, given a new spectrum $\boldsymbol{x}$, the autoencoder is used to reconstruct the spectrum $\hat{\boldsymbol{x}}$. The reconstruction error, which is measured using the Euclidean distance (L2 norm) or Manhattan distance (L1 norm), is used to define the anomaly score:

$$AS_{AE}(\boldsymbol{x}) = ||\boldsymbol{x} - \hat{\boldsymbol{x}}|| \quad or \quad |\boldsymbol{x} - \hat{\boldsymbol{x}}|. \tag{7.3}$$

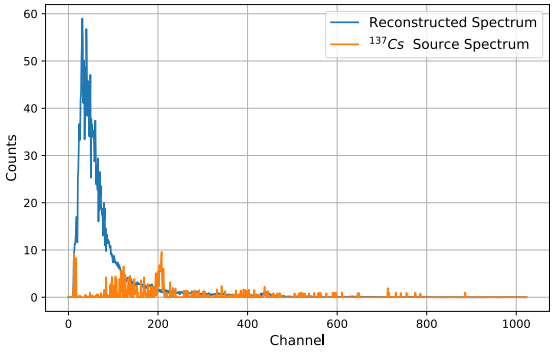
Intuitively, since the autoencoder system is trained using the normal background radiation data, it tends to capture the most significant features of the background data. These captured features can be used to estimate the normal background radiation through reconstruction. When a new background spectrum is given, due to its similarity to the spectra used to train the model, the autoencoder can successfully reconstruct the spectrum without large errors. However, when a spectrum that contains radioactive source information is given, since it has differences from the normal background spectra, the reconstruction error will be large. To better explain this, PCA and autoencoder are applied to a sample dataset consisting of different radioactive sources (^{137}Cs, ^{60}Co, and ^{152}Eu). The dataset was measured in laboratory using Ortec 905 detector for 30 seconds, and the results are shown in Fig. 7.5.

Here, the PCA model is implemented using Scikit-learn [94] and autoencoder models[4] are implemented using TensorFlow [95]. Both models are trained using the measured background radiation data. For autoencoder, batch size is set to be 512, and early stopping [70] is applied to prevent overfitting. New measurements that contain radioactive source information are reconstructed using the PCA and the autoencoder models. These are shown in Fig. 7.5. The original spectra, reconstructed spectra, and the difference (only keeping the positive values) are shown in different colors. Figures 7.5a, 7.5d, and 7.5g contain source information from ^{137}Cs, ^{60}Co, and ^{152}Eu. Figures 7.5b, 7.5e, and 7.5h show results from the PCA. First 65 eigenvectors are used to reconstruct the spectra, which maintain 70% of the total variance. Figures 7.5c, 7.5f, and 7.5i are reconstructed spectra from autoencoder. The results show that both PCA and autoencoder models could correctly capture the source information under current configuration with some noise.

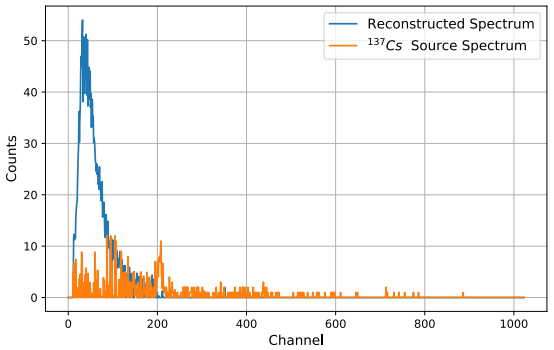
[4]The sample code for autoencoder systems inplemented in TensorFlow is listed in Appendix L.



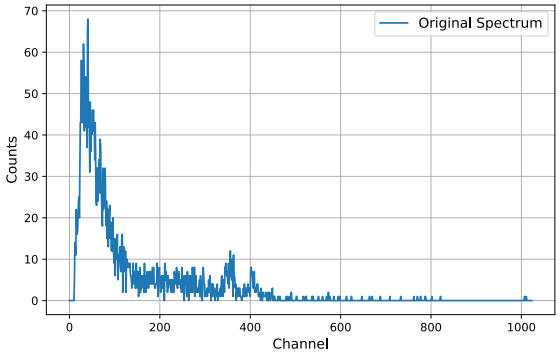
(a) Measured spectrum for ^{137}Cs.



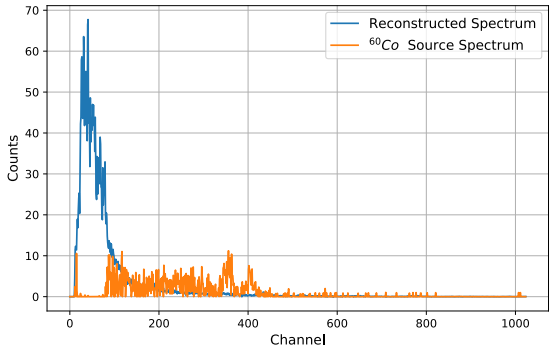
(b) PCA results for ^{137}Cs.



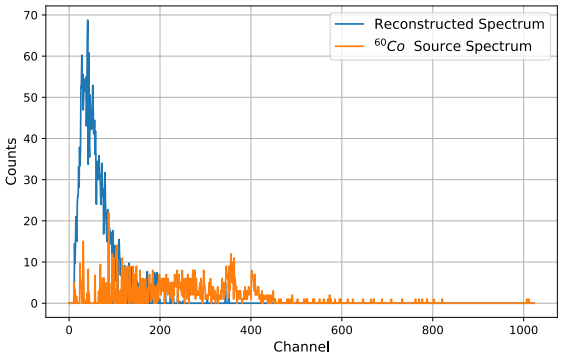
(c) Autoencoder results for ^{137}Cs.



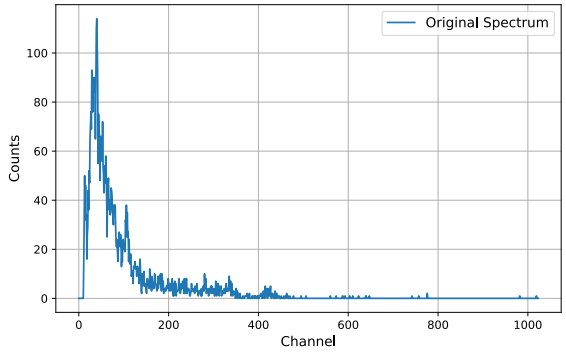
(d) Measured spectrum for ^{60}Co.



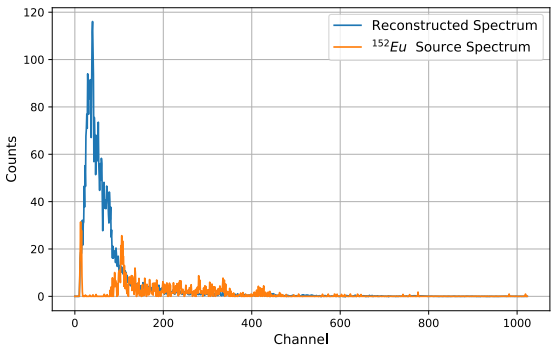
(e) PCA results for ^{60}Co.



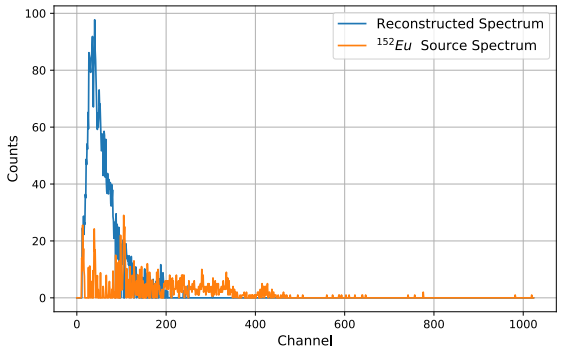
(f) Autoencoder results for ^{60}Co.



(g) Measured spectrum for ^{152}Eu.



(h) PCA results for ^{152}Eu.



(i) Autoencoder results for ^{152}Eu.

Figure 7.5: Spectrum reconstruction examples using PCA and autoencoder. Spectra were collected using 2 × 2 inches NaI detector in laboratory for 30 s.

To show the overall performance of PCA and autoencoder models, PCA and autoencoder are trained and applied using the simulated data which are modeled based on the Ortec 905 detector[5]. Using the anomaly scores defined in Eq. 7.2 and Eq. 7.3, the ROC curves for PCA and autoencoder can be calculated. The results are shown in Fig. 7.6. The PCA model gives the AUC of 0.83 while the autoencoder gives the AUC of 0.89. It is clear that autoencoder outperforms PCA model on the simulated data.

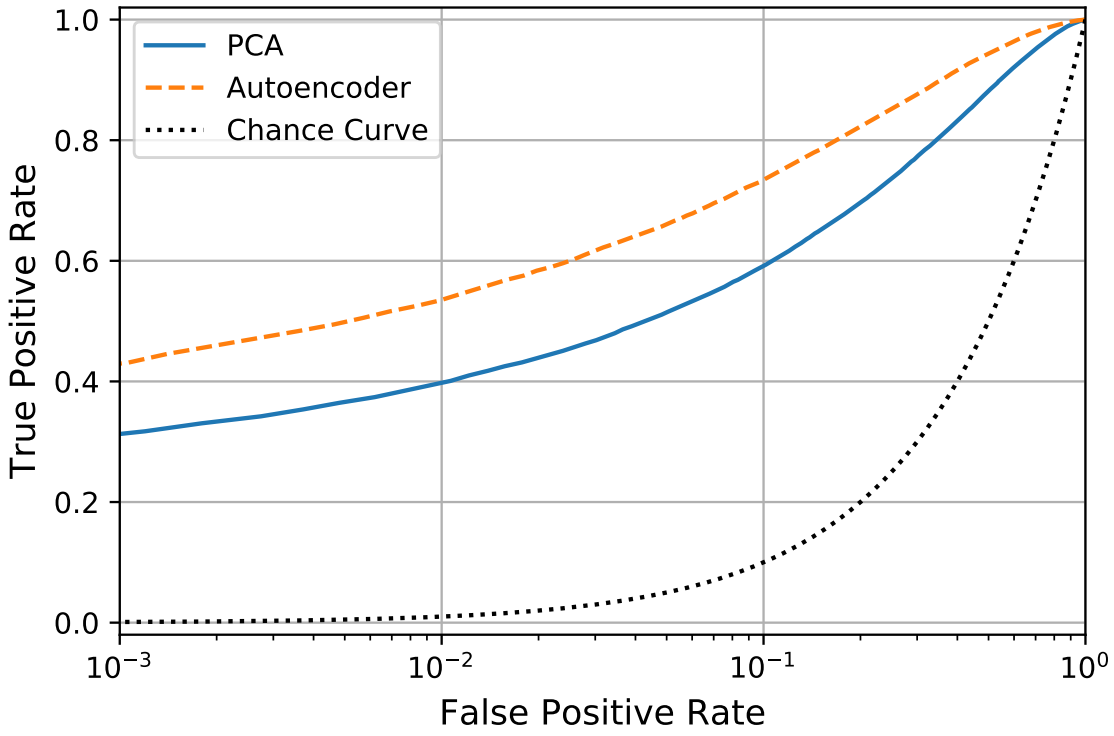


Figure 7.6: ROC curves for PCA and autoencoder systems for anomaly detection.

The structure of the autoencoder system determines its performance for anomaly detection. Two major factors are the number of layers and the number of nodes at each layer. Table 7.1 lists a series of autoencoder structures explored in this work. For example, the structure of 512 - 384 - 256 - 384 - 512 means that the autoencoder has 512 input nodes and 512 output nodes. And there are three hidden layers with 384, 256, and 384 nodes in each layer. Mean square error (MSE) and mean absolute error (MAE) are chosen as the loss function. Accordingly, L2 distance and L1 distance are chosen as the metric to define anomaly score. AUCs for each case are listed in the table. In all the cases, autoencoder with L1 distance metric gives better results compared with the L2 distance metric. Generally, when the number of layers is fixed, changing the

[5]For more details about the simulation, see section 3.2.2.

number of nodes for each layer has little influence on the performance of autoencoder, especially for the cases with L2 distance metric and when the number of layers is 5 or 7. Increasing the number of layers does improve the system's performance, especially for the cases with L1 distance metric and when the number of layers increases from 5 to 7. However, as the number of layers increases, the model's complexity also increases. This will take more iterations to train the model and it is easier to overfit.

Table 7.1: AUC for Different Autoencoder Structures

Number of Layers	**Autoencoder Structure**	**AUC with L2**	**AUC with L1**
3	512 - 384 - 512	0.85	0.91
	512 - 256 - 512	0.85	0.91
	512 - 128 - 512	0.86	0.90
	512 - 64 - 512	0.90	0.91
5	512 - 384 - 256 - 384 - 512	0.86	0.91
	512 - 384 - 128 - 384 - 512	0.89	0.91
	512 - 256 - 128 - 256 - 512	0.89	0.91
	512 - 256 - 64 - 256 - 512	0.90	0.93
7	512 - 384 - 256 - 128 - 256 - 384 - 512	0.89	0.93
	512 - 384 - 256 - 64 - 256 - 384 - 512	0.90	0.94
	512 - 256 - 128 - 64 - 128 - 256 - 512	0.90	0.93
	512 - 384 - 128 - 64 - 128 - 384 - 512	0.91	0.93

7.3 Regression Modeling

PCA and autoencoder assume that the normal background radiation spectrum can be represented in a new space with fewer dimensions. Similarly, since each dimension is usually not independent of other dimensions, another idea is to use the dependency relationship between each dimension to identify the abnormal measurements. In this section, a linear regression model is implemented for anomaly detection.

7.3.1 Methodology

For radiation spectra, each dimension is not totally independent of other dimensions. Thus, one intuitive idea is to predict the information in one dimension using the information from other dimensions. In other words, in regression modeling, for any chosen dimension, it is treated as the dependent variable, and all

the other dimensions are considered as independent variables. A linear regression model is built and the deviation of the dependent variable from the predicted value is used to identify abnormal measurements.

Considering the feature vector $\boldsymbol{x}$ with m dimensions (in this chapter, m equals 512), for each dimension x_i, a linear regression model is built such that:

$$x_i = \sum_{j=1, j\neq i}^{m} \beta_{ij} \cdot x_j + \beta_{i0} + \epsilon_i, \tag{7.4}$$

where i represents the dependent dimension to be regressed, j represents the independent variable direction, β_{ij} are the coefficients, and ϵ_i is the deviation.

The linear regression model is fitted on n instances $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(n)}$ and the coefficient β_{ij} are determined through minimizing the mean square error (MSE). In this way, m different linear models are built for each dimension from 1 through m. For the new measurement $\boldsymbol{x} = (x_1, x_2, \cdots, x_m)$, the m linear models are applied and deviations $\epsilon_1, \epsilon_2, \cdots, \epsilon_m$ are calculated. To measure the anomaly score, two commonly used metrics are mean squared error (MSE) and mean absolute error (MAE):

$$MSE(\boldsymbol{x}) = \frac{1}{m} \sum_{j=1}^{m} \epsilon_j^2 \tag{7.5}$$

$$MAE(\boldsymbol{x}) = \frac{1}{m} \sum_{j=1}^{m} |\epsilon_j|. \tag{7.6}$$

The above methods assume that the contribution of different dimensions are equally important. In real cases, some dimensions are more important in determining abnormal measurements while other dimensions are less important. For example, to determine the existence of ^{60}Co, the counts at channels corresponding to energies 1.17 MeV and 1.33 MeV are more important than other channels. To incorporate this, in the so-called attribute-wise leaning for scoring outliers (ALSO) approach, Paulheim and Meusel [96] proposed weighted score to measure the anomaly score.

Considering n instances $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(n)}$, for the i-th dimension, define the root relative squared error R_i as:

$$R_i = \sqrt{\frac{\sum_{j=1}^{n} \epsilon_i^{(j)}}{\sum_{j=1}^{n} (x_i^{(j)} - \bar{x}_i)^2}}, \tag{7.7}$$

where $\bar{x}_i$ is the average value for the i-th dimension. Then, the corresponding weight w_i for the i-th dimension is defined as:

$$w_i = 1 - \min(1, R_i). \tag{7.8}$$

Intuitively, the weight w_i is 0 if the model performs worse then simply predicting average. Then, this dimension is assumed to be an irrelevant dimension for anomaly detection [16]. Thus, the outlier score for linear regression model is defined as:

$$score(\boldsymbol{x}) = \sum_{i=1}^{m} w_i \cdot \epsilon_i^2(\boldsymbol{x}). \tag{7.9}$$

7.3.2 Implementation

When implementing linear regression model for anomaly detection, certain factors need to be considered, including semi-supervised or unsupervised cases, and to normalize features or not. For semi-supervised cases, it is assumed that the normal background radiation data are available. Then the linear regression model is fitted based on the normal measurements and the fitted model is used to make predictions for new measurements. For unsupervised cases, the linear regression model is fitted based on all available data. Feature normalization refers to the process of subtracting the mean value from input data and then dividing the standard deviation for each dimension. Through feature normalization, the transformed input data have zero-mean and unit variance for each dimension.

In this section, for the radiation data, AUC is calculated using MSE (Eq. 7.5), MAE (Eq. 7.6), and weighted score (Eq. 7.9) as anomaly score. The results are shown in Table 7.2. The configuration, feature normalization, and metric have different influences. Compared with MSE, MAE always gives better results. And in the unsupervised implementation, after feature normalization, weighted score performs much better than using MSE and MAE. The unsupervised implementation method using MAE as anomaly score and without feature normalization gives best result among all the configurations.

Table 7.2: Comparison of Anomaly Detection for Linear Regression

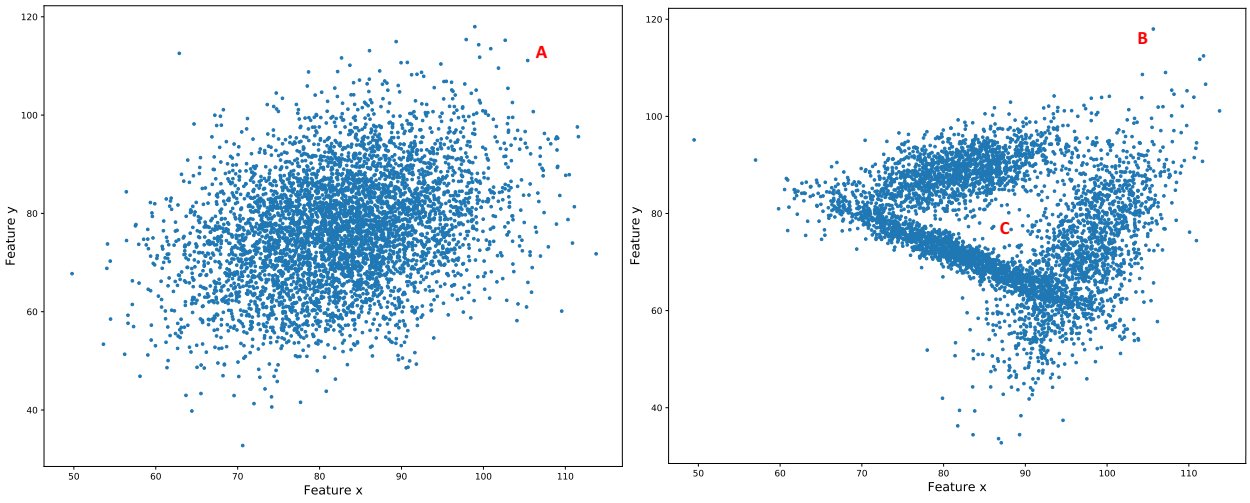
Configuration	**Semi-Supervised**			**Unsupervised**		
	MSE	**MAE**	**Weighted Score**	**MSE**	**MAE**	**Weighted Score**
No Feature Normalization	0.9104	0.9216	0.8974	0.9065	0.9306	0.8932
Feature Normalization	0.8274	0.9016	0.8510	0.7338	0.8866	0.9289

7.3.3 Discussion

This section discusses the application of feature-wise linear regression for anomaly detection, and the results are compared for different configurations. It is worth mentioning that the principles behind the above models are not limited to using a linear regression model as the base model. Other models, such as regression trees [97] or even more advanced random forest models, can also be used as the base model, which might provide more robust performance for anomaly detection.

7.4 Proximity-Based Models

Previous sections discussed the application of PCA, autoencoder, and regression model for anomaly detection. PCA and linear regression models work well for the cases where all the normal background data are similar to each other. In other words, in the feature space, these normal background radiation spectra form one "cluster" as shown in Fig. 7.7a. In the figure[6], features x and y represent two general features. The feature space in practical cases is usually more than two dimensions. For nuclear radiation, feature x and feature y can be considered as the measured counts at channels corresponds to energies 1.17 MeV and 1.33 MeV, which correspond to the peak energies of ^{60}Co. In this case, the data within the cluster correspond to normal background radiation data. Outliers such as measurement A can be thought of as the spectrum containing information from ^{60}Co, thus defined as abnormal measurement.



(a) Anomaly detection with one cluster.

(b) Anomaly detection with three clusters.

Figure 7.7: Illustration of Anomaly Detection. Sample abnormal measurements are denoted by characters A, B, and C.

[6]All data are just used to illustrated the idea behind proximity-based method. There is no connection with actual radiation data.

In addition to the case shown in Fig. 7.7a, however, there are cases where there might exist multiple dense clusters. For example, consider the case shown in Fig. 7.7b, in which there are three dense clusters. In this case, PCA and linear regression models might still work for abnormal measurement B, but they cannot correctly identify measurement C as abnormal measurement without appropriate feature transformation. For abnormal measurements like B, methods based on proximity (distance or density) would be appropriate. In this section, proximity-based anomaly detection techniques are studied. More specifically, distance-based and density-based methods are discussed.

7.4.1 Distance Metrics

Distance, or dissimilarity, is the metric used to measure the dissimilarity between measurements. In many cases, especially for proximity-based models, the anomaly detection process finally becomes the process of measuring the distance between the new measurement and calculated template or other measurements.

Well-defined distance metrics generally have some desired properties, including positivity, symmetry, and triangle inequality [98]. Suppose $\boldsymbol{x}$ and $\boldsymbol{y}$ are two measurements with m dimensions, and $d(\boldsymbol{x}, \boldsymbol{y})$ is the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, then the following properties should hold for the distance metric $d(\cdot)$:

1. Positivity
   - $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$ for all $\boldsymbol{x}$ and $\boldsymbol{y}$.
   - $d(\boldsymbol{x}, \boldsymbol{y}) = 0$ only if $\boldsymbol{x} = \boldsymbol{y}$.
2. Symmetry

   $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$ for all $\boldsymbol{x}$ and $\boldsymbol{y}$.
3. Triangle Inequality

   $d(\boldsymbol{x}, \boldsymbol{z}) \leq d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z})$ for all $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$.

There are several commonly used distance metrics. Expressions for Euclidean distance, Manhattan distance, Mahalanobis distance, and Cosine distance are:

- Euclidean distance
$$d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{m} (x_i - y_i)^2}. \tag{7.10}$$
- Manhattan distance
$$d(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{m} |x_i - y_i|. \tag{7.11}$$

- Mahalanobis distance

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(\boldsymbol{x} - \boldsymbol{y})^T S^{-1} (\boldsymbol{x} - \boldsymbol{y})}, \tag{7.12}$$

  where S is the covariance matrix.

- Cosine distance

$$d(\boldsymbol{x}, \boldsymbol{y}) = 1 - \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{||\boldsymbol{x}|| \; ||\boldsymbol{y}||}. \tag{7.13}$$

Additionally, Euclidean distance and Manhattan distance are two special cases of Minkowski distance. The more general expression of Minkowski distance is:

$$d(\boldsymbol{x}, \boldsymbol{y}) = \left(\sum_{i=1}^{m} (x_i - y_i)^r \right)^{1/r}. \tag{7.14}$$

In addition to above distances, there are many other distances, such as Hamming distance and Jaccard dissimilarity. As will be shown later, choosing the correct distance metrics is important and the performance of anomaly detection can be influenced by the choice of distance metrics.

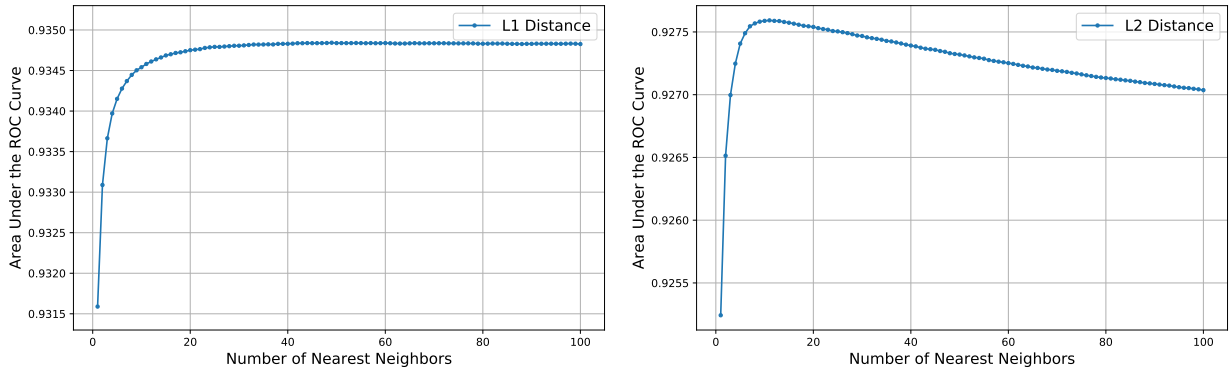
7.4.2 Distance-Based Anomaly Detection

For distance-based method, the intuitive idea is to compare the distance to its nearest neighbors to identify the abnormal measurements, which leads to k-nearest neighbors (KNN) method. Considering measurements A, B, and C in Fig. 7.7, they share one common characteristic: compared with majority of measurements, they are located in a relatively sparse area in the feature space. In other words, they are far away from their neighbors.

To formalize the above idea, suppose that there are n normal measurements $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(n)}$, and the new measurement to be considered is $\boldsymbol{x}$. Among the existing measurements, without loss of generality, suppose that the first k closest measurements, or nearest neighbors, are denoted by $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(k)}$. Then the anomaly score can be defined using the distance to these k measurements. Commonly used anomaly scores include the distance to the k-th nearest neighbor, average distance to the first k nearest neighbors, and the Harmonic distance to the first k nearest neighbors [16].

The performance of k-nearest neighbors method is mostly influenced by the number of nearest neighbors to be used, the definition of distance metric, and the definition of anomaly score. Simulation study shows that using the definition of average distance to the k nearest neighbors and the Harmonic distance to the k nearest neighbors have very similar results, which are both better than using the distance to the k-th nearest neighbor. To show the influence of number of neighbors on the performance of KNN method, the

Manhattan distance (L1 norm) and Euclidean distance (L2 norm) are implemented, and the anomaly score is calculated using the average distance to the k nearest neighbors. The results are shown in Fig. 7.8.



(a) KNN with Manhattan distance (L1).

(b) KNN with Euclidean distance (L2).

Figure 7.8: Comparison of KNN with L1 and L2 distances for different number of nearest neighbors.

In Fig. 7.8, using 49 neighbors, the AUC when using Manhattan distance can reach 0.9348, while for Euclidean distance, with 11 neighbors, the AUC is 0.9276. In addition, Fig. 7.8 clearly shows that Manhattan distance (L1) are less sensitive to number of neighbors than Euclidean distance (L2) when the number of nearest neighbors is large enough. This is due to the fact that Manhattan distance is less sensitive to outliers, which makes it a more robust method for anomaly detection.

Rectified Distance

Choosing the appropriate distance metric is the key to KNN method. For different problems, some domain-specific distances or pseudo-distances can be defined to improve the performance of anomaly detection systems. Here, considering the properties of radiation spectrum, the so-called rectified distances are proposed in this work for semi-supervised anomaly detection based on traditional Euclidean distance and Manhattan distance:

- Rectified Euclidean distance

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{m} (\max(x_i - y_i, 0))^2} \tag{7.15}$$

- Rectified Manhattan distance

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{m} \max(x_i - y_i, 0), \tag{7.16}$$

  where $\boldsymbol{x}$ is the new measurement, $\boldsymbol{y}$ is the normal measurements (for example, normal background radiation spectrum) from the training dataset.

Essentially, for semi-supervised anomaly detection, if the new measurement has smaller value for some dimensions compared with the normal measurements, in rectified distance, these dimension are not important at all. Clearly, the definition of rectified distance does not follow the three properties described in section 7.4.1. But this does not influence the practical implementation of rectified distances.

In the comparison experiment, it is found that using the rectified Euclidean distance, AUC is increased from 0.927 to 0.936, while for rectified Manhattan distance, AUC is increased from 0.934 to 0.940. Though the impact is small, both suggest that the rectified distances are slightly better choices.

7.4.3 Density-Based Anomaly Detection

K-nearest neighbors method is based on comparing the distance to its nearest neighbors. It works well when different clusters have similar densities (as shown in Fig. 7.7b). However, when the density difference among different clusters is large, the performance of k-nearest neighbors method is significantly influenced. Figure 7.9 shows a sample case where there are two clusters with different densities[7]. For this case, KNN method can easily identify points like B as abnormal measurements since it is far away from its neighbors, but it might fail to identify point A as an abnormal measurement since it is close to a dense cluster.

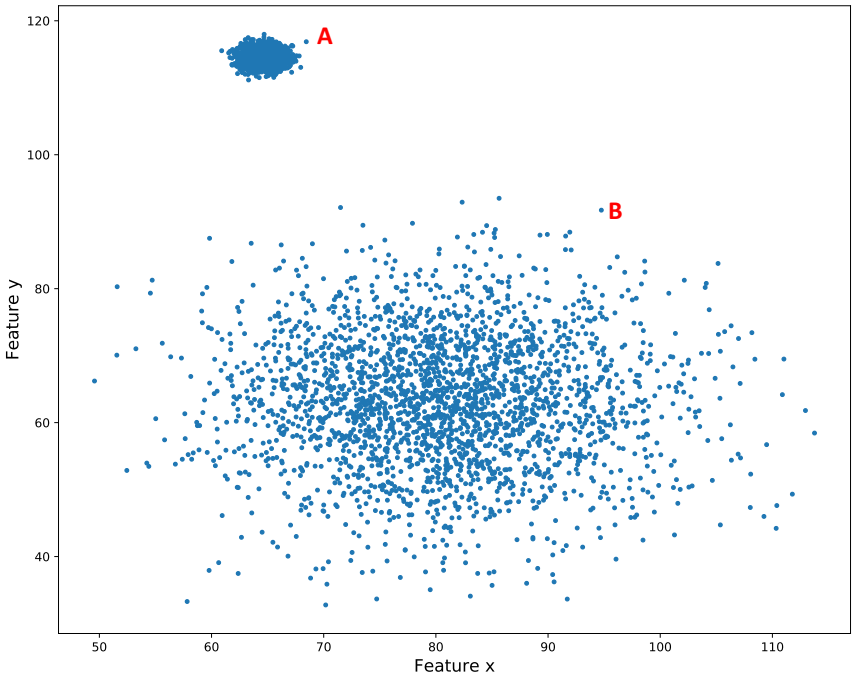


Figure 7.9: Illustration of Anomaly Detection. Two sample abnormal measurements are denoted by letters A and B.

[7]All data are just used to illustrated the idea behind density-based method. There is no connection with actual radiation data.

To address this challenge, the measurement of local density is used for anomaly detection. Methods such as local outlier factor (LOF) [99] have been developed. For the radiation data used in this chapter, LOF has similar performance as KNN method on the simulated data.

7.4.4 Discussion

Although distance-based and density-based methods work for both semi-supervised and unsupervised cases, in practical implementation, the computational complexity needs to be considered and unsupervised implementation is not suggested. Appendix M includes the analysis on time complexity of KNN method. As shown in Fig. M.2, as the sample size increases, the query time increases linearly. For unsupervised cases, the test size is usually much larger compared with the training set in semi-supervised implementation, which will dramatically increase the query time.

For the high dimension problems, due to *the curse of dimensionality*, the distances between measurements tend to increase and it is more difficult to find "close" neighbors to make predictions. However, the distance-based and density-based methods actually work for radiation spectrum regardless of high dimensionality problems. The possible reason is that although the spectrum has 512 dimensions, most useful information is concentrated in the first 100 to 200 dimensions, which is clearly shown in Fig. 7.1. Thus, the actual dimensionality of the problem is much smaller, which makes the anomaly detection techniques based on distance or density still applicable for nuclear radiation spectra.

7.5 Summary

In this chapter, several commonly used anomaly detection techniques are discussed in detail. Simulation data are used to quantify and compare the performance of different algorithms. Table 7.3 lists the performance of studied algorithms for spectrum-based anomaly detection.

Table 7.3: Area Under the ROC Curve (AUC) of Different Anomaly Detection Techniques

Method	**Semi-supervised Implementation**	**Unsupervised Implementation**
PCA	0.831	N.A.
Autoencoder	0.936	N.A.
Linear Regression	0.922	0.931
KNN	0.940	N.A.

For the most simple method, the PCA method, although its performance (AUC 0.831) is poor compared

with other methods, it is easy and fast to build the model and make predictions. For the autoencoder system, its performance (AUC 0.936) is much better compared with the PCA method and it is also fast to make predictions, but it is harder to train the model, especially when the model is too complex. For the linear regression model, the unsupervised implementation (AUC 0.931) gives better performance than the semi-supervised implementation (AUC 0.922). For the KNN method, it is very easy to understand and gives the highest AUC (0.940) among these methods. However, as a lazy method, the KNN method may take longer computational time to make predictions, especially when the training dataset is too large. In addition to the methods discussed in this chapter, other methods, such as one-class support vector machine (SVM), methods based on clustering analysis, as well as ensemble methods such as isolation forest [100, 101], can also be applied.

As one important step after data collection, anomaly detection provides important prior information for source localization and isotope identification. Accurately identifying abnormal measurements can reduce the computation resources, restrict the searching region for source localization, and simplify the isotope identification tasks[8].

[8] A preliminary discussion of potential usage of anomaly detection for isotope identification is included in Appendix N.

Chapter 8

Summary and Future Work

8.1 Summary

Nuclear radiation detection plays an important role in public health and national security. Nuclear radiation detection process consists of data collection, anomaly detection, and source localization procedures. After nuclear radiation data are collected, anomaly detection can be first applied and abnormal measurements are identified. Through anomaly detection, the volume of data to be further processed is dramatically reduced and suspicious regions that might contain radioactive sources are identified. Then, source localization techniques would be applied to pinpoint the radioactive sources based on the result from anomaly detection. This work studied the application of mobile sensor networks for nuclear radiation detection in detail.

More specifically, in this work, a mobile sensor network simulation system based on GADRAS was first developed. The simulation system could simulate multiple mobile detectors in a simplified urban environment. GADRAS simulated data made the simulation more realistic and the simulated data were not restricted to count rate data anymore. The simulation system provided the foundation for this project. In addition to simulations, a small mobile sensor network system was also built with D3S detectors and the simulated data injected in measured background radiation data were obtained to test the performance of proposed algorithms in near-real environment.

A systematic analysis on the performance of mobile sensor networks was performed based on simulations. Using the radiation count rate data, the influence of anomaly detection algorithms, detector characteristics, and number of detectors were analyzed in terms of anomaly detection, SNR and so on. These analyses would be helpful for the actual deployment of mobile sensor networks in cities.

For anomaly detection, when background radiation levels were easy to obtain, the traditional k-sigma method was the first choice. It was efficient and easy to implement. When background radiation levels were not available, inspired by geospatial analysis, this work proposed the Poisson kriging model for radiation count data. The geospatial correlation between radiation measurements was explicitly considered and the radiation levels were estimated at unknown locations using available measurements. An anomalous radioac-

tive source detection method was then proposed based on SNR. Results showed that the proposed algorithm could correctly estimate the background radiation distribution and reach over 90% accuracy for radioactive source detection when source was close to detectors (e.g., less than 5 meters away) or source intensity was high enough (e.g., 300 μCi or above).

Poisson kriging was helpful for identifying suspicious regions that might contain radioactive sources. Based on the results from Poisson kriging, two source localization techniques, KDE with MLE and grid search with MLE, were proposed to estimate the locations of radioactive sources. Simulation studies and experiments with simulated radioactive source injected compared the performance of the proposed algorithms and showed that grid search with MLE gave robust estimation and could reduce the source localization error to be within 3 meters.

Lastly, this work conducted an exploratory study on spectrum-based anomaly detection for radiation data. Various methods, including methods based on dimensionality reduction, methods based on regression analysis, and methods based on proximity, were studied in detail. The performance of different methods was compared using simulated data. Methods such as the KNN method or the autoencoder method, performed well for anomaly detection, but they had limitations on prediction complexity or training complexity. Methods such as the PCA model might not give the best performance, but they were easy to implement or make prediction. Trade-off between model complexity and model performance need to be considered for actual implementation.

Overall, this work presented the workflow behind nuclear radiation detection and demonstrated the feasibility of mobile sensor networks for nuclear radiation detection. Through simulations and experiments, the performance and limitation of the proposed approaches were quantified. These results and analyses would provide good reference for the actual deployment of mobile sensor networks and future related work.

8.2 Future Work

As a challenging problem, nuclear radiation detection consists of many different aspects. This work only covers parts of this complex problem. Influence of other factors such as radiation temporal variation, mobile radioactive sources, and shielding effect have not been explicitly studied. For future work, there are several potential research directions:

Implementing and testing the proposed algorithms in real-world environments. The algorithms proposed in this work have been tested through simulations and experiments. However, there are certain simplifications for the simulation, and the experiment was conducted only in a small region without

actual radioactive sources. To obtain realistic assessment of the proposed algorithms, these algorithms need to be implemented and tested in real-world environments.

Combining information from different detectors. This work assumes that all data are collected using the same types of detectors. This assumption simplifies the calculation of Poisson kriging and MLE. However, in real cases, information from different types of detectors may need to be combined. For example, there exists different types of detectors, which may have different resolution, efficiency, and channel numbers. In this case, developing efficient algorithms that can combine information from different detectors might be necessary.

Combining stationary and mobile sensor networks. Although this work focuses on mobile sensor networks, stationary radiation detectors are also very important. Mobility brings certain advantages for nuclear radiation detection, but it also brings restrictions such as detector size. Stationary detectors, on the other hand, can be placed at choke points with more sensitive detectors. In this case, it is necessary to develop algorithms that can combine information from stationary and mobile detectors for nuclear radiation detection.

Hardware development. As shown in this work, the number of detectors is a key factor that determines the performance of mobile sensor networks. More detectors can help increase the probability of radioactive sources being detected and increase the source localization accuracy. However, the cost restricts the deployment of more detectors. Development of cheaper radiation detectors with similar performance makes it possible to deploy more detectors and achieve better performance.

Information fusion for object tracking. This work only studies the application of radiation detectors for radiation detection and monitoring. In real cases, the information from other sources, for example, surveillance cameras, can also be utilized. For example, through combining radiation detection techniques and computer vision techniques such as convolutional neural networks, the movement of vehicles that carry radioactive sources can be tracked in cities, which is important for detecting and tracking mobile radioactive sources.

Appendix A

Detector Characteristics

In this work, simulations are conducted based on two types of detectors, the Ortec 905-3 NaI detector and the D3S detector.

The Ortec 905-3 detector, is a NaI(Tl) scintillation detector from ORTEC©. It comes with $2'' \times 2''$ crystal and $2''$ PMT. The dimension of the 905-3 detector is shown in Fig. A.1.

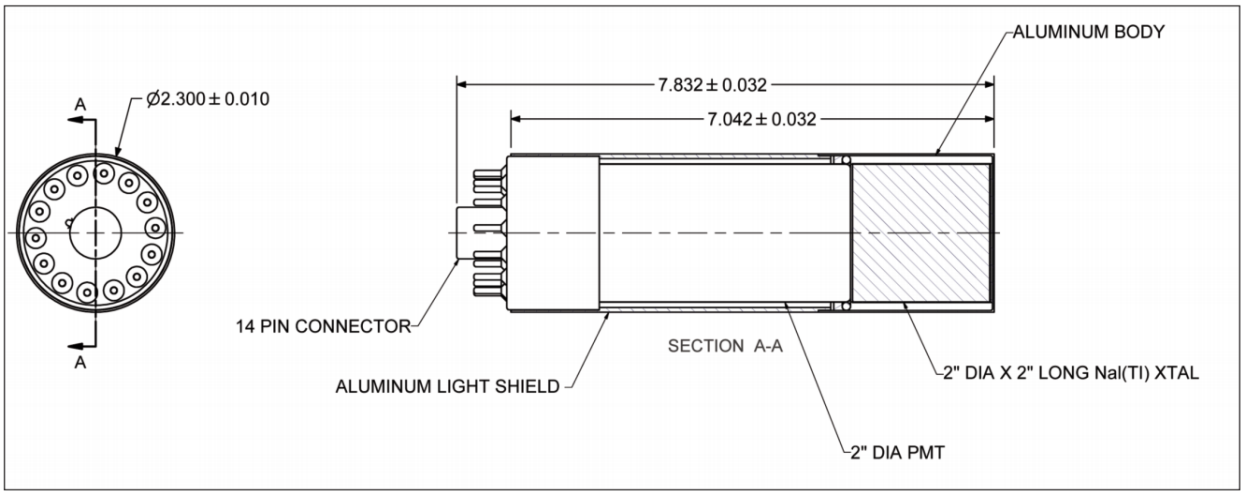


Figure A.1: Dimensions of NaI 905-3 detector.

More detailed information about the Ortec 905-3 detector can be found at `http://www.ortec-online.com/-/media/ametekortec/brochures/905.pdf`.

Another detector, the Discreet Dual Detector, or D3S, is a hybrid gamma/neutron detector used for national security applications from Kromek Group plc. The detailed technical information is shown in Table A.1.

Table A.1: Characteristics of D3S Detector

Item	**Specifications**
Gamma detector	CsI(Tl) detector
Resolution	7% resolution at 662 keV
Gamma energy range	30 keV to 3 MeV
Gamma sensitivity	500 cps/μSv/h (5 cps/μR/h) for Cs137
Maximum throughput for gamma channel	$10,000$ *cps*
Dose rate	Up to 15 μSv/h at 662 keV (1.5 mR/h)
Neutron detector	Sensitivity 12 *cps*/nv
Neutron detector gamma rejection	Better than 10^{-7}
Maximum throughput for neutron channel	$10,000$ *cps*
Communications	Micro USB Bluetooth$^{@}$
Operational battery life	12 hours
Operational temperature range	-20 to 50°C
Size	132 mm $\times$ 80 mm $\times$ 23.5 mm ($5.2^{''} \times 3.1^{''} \times 0.9^{''}$)
Humidity	Up to 93% RH
Moisture/Dust	IP53
Weight	237 g (0.52 lbs)
Battery	1450 mAh Lithium polymer
Charging	Inductive charging Charging via USB
External LED's	Visual detector status

Appendix B

Introduction to GADRAS

Gamma Detector Response and Analysis Software, or GADRAS, is a spectrum simulation and analysis software developed and maintained by Sandia National Laboratories. For radiation spectrum simulation, GADRAS applies a detector response function to compute the output from corresponding detectors [77,102]. Current response function applies a combination of first-principles calculations and empirical models, which makes the simulated spectrum more accurate for a broad range of detectors. For background radiation simulation, it provides templates based on locations, such as the templates from Chicago and Salt Lake City. For isotope identification, GADRAS uses a template-based approach for making identifications [76, 79].

A snapshot of the GADRAS screen is shown in Fig. B.1. The left part of the figure is GADRAS user interface. Users can specify a series of simulation settings, such as the distance to source, detector height, simulation time, background radiation setting and so on. The right part of the figure shows a simulated spectrum of a 10 μCi ^{137}Cs radioactive source for 1 hour.

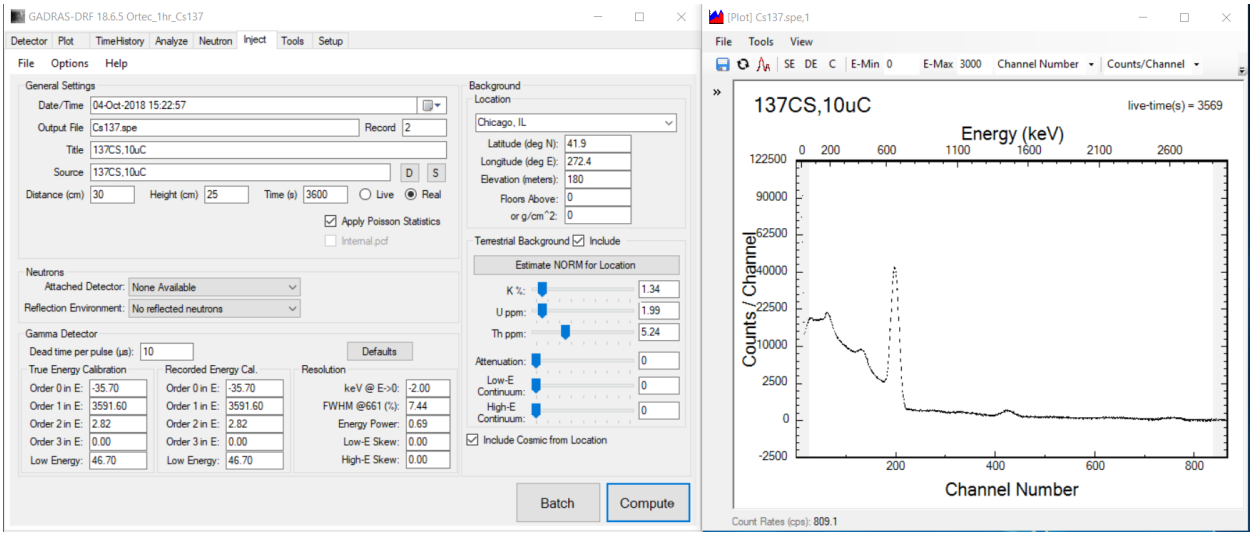


Figure B.1: Snapshot of GADRAS software.

Appendix C

Synthetic Data Simulation Process

The synthetic data simulation includes two parts: background radiation simulation and radioactive source simulation.

For background radiation simulation, to incorporate the spatial variation, high background radiation and low background radiation are needed. In GADRAS, the background radiation simulation can be specified using cities. Since Chicago has relatively low background radiation compared with Salt Lake City, the radiation data simulated in Chicago are used to represent low background radiation, and the radiation data simulated in Salt Lake City are used to represent high background radiation. In both cases, detectors are assumed 1 meter above the ground and spatial location (Chicago or Salt Lake City) is the only changing parameter. Based on these settings, for D3S detector, the average background radiation count rate for Chicago is 31 *cps* while the average background radiation for Salt Lake City is 38 *cps*.

For radiation from radioactive sources, spatial location of radioactive source is no longer the changing parameter. Instead, radioactive isotopes, source strength, distance from the source, and shielding effect are the most influencing factors. In this work, since all data are assumed to be collected in outdoor environment and simulated sources are not strong, the shielding effect except the shielding from air is ignored. During the simulation, radioactive sources are assumed to be placed on the ground and detectors are located 1 meter above the ground. Simulations are conducted for different isotopes, such as ^{137}Cs, ^{60}Co, and ^{54}Mn, and different source strength, such as $100\mu Ci$, $300\mu Ci$, $500\mu Ci$, $1,000\mu Ci$, and $2,000\mu Ci$. For the given isotope and source strength, the simulation is conducted for different distances. The schematic diagram of the simulation process is shown in Fig. C.1. The simulated radioactive source is placed at the center of circles. For each distance from 0.5 m until 20 m, the simulation is conducted for every 0.5 m. At each distance, 1,000 or 3,000 simulated spectra are generated and saved for later sampling process.

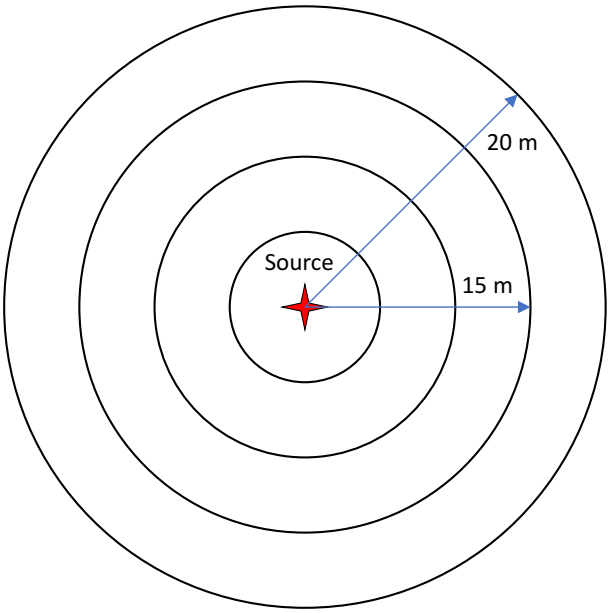


Figure C.1: Schematic diagram of radioactive source spectrum simulation process. The spectra are simulated at different distances from the radioactive source.

Appendix D

Instructions on How to Use the Synthetic Data

D.1 Synthetic Data Simulation

In this work, the spectra from background radiation and different isotopes are simulated using GADRAS as described in section 3.2.2. The background radiation and the radiation from radioactive sources are separated.

The following code demonstrates how to load the simulated background radiation data. The loaded data are saved in Numpy array format such that each row corresponds to one single spectrum.

```
import numpy as np

# load simulated background spectra
background = np.load('./D3_background.npy')
```

Listing D.1: Demo of reading simulated background radiation spectra

For the radiation from radioactive sources, the distance from the source need to be considered. In the following code, the loaded data are in Python dictionary format. The key is the distance from the radioactive source, and the corresponding value is Numpy array which contains multiple simulated spectra.

```
import pickle

# load simulated radioactive source spectra
with open('./D3_source.p', 'rb') as f:
    source = pickle.load(f)
```

Listing D.2: Demo of reading simulated radioactive source spectra

D.2 Mobile Sensor Network Simulation

For the mobile sensor network simulation system, users can specify the region of the simulation, source position, number of detectors, and the speed of the detectors. In the following code, the region are assumed to contain 36 (6×6) $100m \times 80m$ city blocks as shown in Fig. 3.4. There are 100 mobile detectors and the simulation lasts for 3,600 seconds (1 hour).

```
from road_walk_simulator_shielding import road_walk_simulator

# define source locations
source = [(105, 85), (105, 265), (105, 445), (215, 175), (215, 355)]

# build simulator
simulator = road_walk_simulator(area=(6, 6),
                                source_loc=source,
                                detector_num=100,
                                speed_low=1.4,
                                speed_high=1.8,
                                n_channel=512,
                                period=3600,
                                high_back_radius=400,
                                hasSource=False,
                                shielding=False)
# run simulation
simulator.simulate()

# save data to local folder
simulator.save_data('./simulation.csv', spectrum=True)
```

Listing D.3: Demo of mobile sensor network simulation

The above code saves the simulation results as a CSV file. The CSV file contains the information about detector id, timestamps, detector localizations, distance to the nearest detector, count rate information, and the whole spectrum. Figure D.1 shows a snapshot of the CSV file generated from the simulation.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	id	period	x	y	distance	shielding	back_counts	source_counts	gross_counts	s1	s2	s3	s4	s5	s6	s7	s8
2	1	1	547.543762	358.065216	86.9719772	0	31	0	31	0	0	0	0	0	0	0	0
3	1	2	547.543762	359.565216	85.4726334	0	29	0	29	0	0	0	0	0	0	0	0
4	1	3	547.543762	361.065216	83.9733124	0	28	0	28	0	0	0	0	0	0	0	1
5	1	4	547.543762	362.565216	82.4740143	0	38	0	38	0	0	0	0	0	0	0	0
6	1	5	547.543762	364.065216	80.9747391	0	22	0	22	0	0	0	0	0	0	0	1
7	1	6	547.543762	365.565216	79.4754944	0	34	0	34	0	0	0	0	0	0	0	0
8	1	7	547.543762	367.065216	77.9762726	0	25	0	25	0	0	0	0	0	0	0	1
9	1	8	547.543762	368.565216	76.4770889	0	45	0	45	0	0	0	0	0	0	0	3
10	1	9	547.543762	370.065216	74.9779358	0	41	0	41	0	0	0	0	0	0	0	0
11	1	10	547.543762	371.565216	73.4788132	0	26	0	26	0	0	0	0	0	0	0	0
12	1	11	547.543762	373.065216	71.9797363	0	26	0	26	0	0	0	0	0	0	0	0
13	1	12	547.543762	374.565216	70.48069	0	31	0	31	0	0	0	0	0	0	0	1
14	1	13	547.543762	376.065216	68.9816895	0	25	0	25	0	0	0	0	0	0	0	0
15	1	14	547.543762	377.565216	67.4827347	0	33	0	33	0	0	0	0	0	0	0	0
16	1	15	547.543762	379.065216	65.9838257	0	27	0	27	0	0	0	0	0	0	0	1
17	1	16	547.543762	380.565216	64.4849625	0	30	0	30	0	0	0	0	0	0	0	0
18	1	17	547.543762	382.065216	62.9861603	0	33	0	33	0	0	0	0	0	0	0	1
19	1	18	547.543762	383.565216	61.4874115	0	28	0	28	0	0	0	0	0	0	1	0
20	1	19	547.543762	385.065216	59.9887276	0	28	0	28	0	0	0	0	0	0	1	1
21	1	20	547.543762	386.565216	58.4901123	0	15	0	15	0	0	0	0	0	0	0	0
22	1	21	547.543762	388.065216	56.9915695	0	33	0	33	0	0	0	0	0	0	1	2
23	1	22	547.543762	389.565216	55.493103	0	36	0	36	0	0	0	0	0	0	0	1
24	1	23	547.543762	391.065216	53.9947243	0	35	0	35	0	0	0	0	0	0	0	1

Figure D.1: Snapshot of simulated radiation data.

Appendix E

Simulation Procedures

In this work, the simulation is conducted in a region segmented into a 6 × 6 grid of 36 blocks as shown in Fig. E.1. Section 3.3 gave a brief description of the simulation. This chapter gives more details about the simulation, especially for the cases corresponding to Chapter 4.

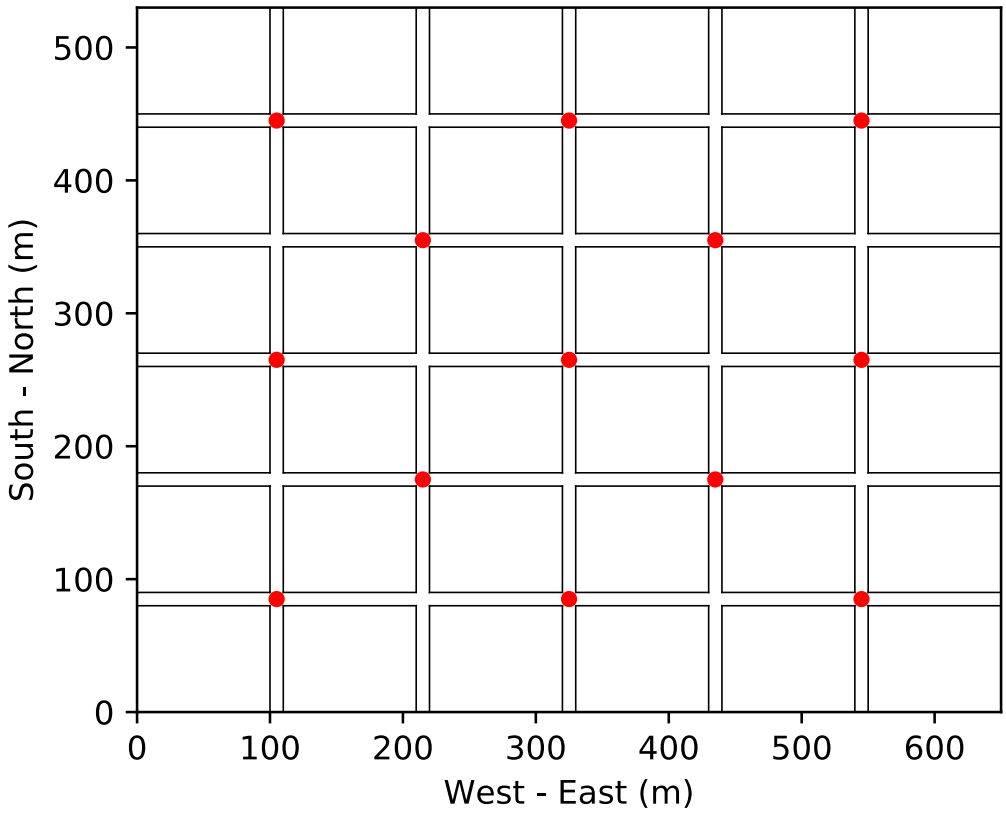


Figure E.1: Illustration of simulated radioactive source locations.

Consider the case shown in Fig. E.1, 13 radioactive source are placed at the intersections of roads[1], denoted by red circles. These sources cover different cases, and they are far away from each other such that for any detector in this region, the measured radiation data is influenced by at most one radioactive source.

[1]The exact locations of radioactive sources have little influence on the conclusion.

GADRAS simulated data as described in section 3.2.2 contain two parts: background radiation data and radioactive source data. For background radiation data, the data simulated based on Chicago is used to model low background radiation region and the data simulated based on Salt Lake City is used to model high background radiation region. For radioactive source data, at each distance from 0.5 m to 20 m, there are 1,000 or 3,000 spectra simulated for sampling purpose. Appendix C and Appendix D have more details about the simulated data.

After the simulation starts, multiple radiation detectors are moving in the simulated area. When there are no radioactive sources or the detector is far away from the radioactive sources, the count rate from radioactive source is set to be 0. The background radiation count rate is sampled from the simulated background radiation database. In this work, since the lower left region is set to be high background region, when the detector moves into the high background radiation region, the background radiation data is sampled from the data simulated based on Salt Lake City, and when the detector moves into the low background radiation region, the background radiation data is sampled from the data simulated based on Chicago. When there are radioactive sources, such as the cases in Fig. E.1, the background radiation data is still sampled as described above. For source data, at first, the distances of the detector to all radioactive sources are calculated. If the smallest distance is larger than 20 m, the detector are not influenced by the radioactive source and the source count rate is set to be 0. When the smallest distance is smaller than 20 m, the source data is sampled from the simulated radioactive source data according to the smallest distance and the source count rate is larger than 0. This process is repeated for every second during the simulation.

Appendix F

Experimental Procedures

In this work, the experiment with a small mobile sensor network system was conducted on the engineering campus of the University of Illinois at Urbana-Champaign. The experiment was conducted in an outdoor region of approximately 500 × 400 m^2 in area. During the experiment, operators placed the D3S detectors in their pockets and walked at the normal walking speed (around 1.4 m/s) along the designed paths. The experimental region and walking paths are shown in Fig. F.1. There are several regions that have relatively higher background radiation. These locations are all near buildings. Three typical high background radiation locations are denoted by dashed rectangles, including Nuclear Radiation Laboratory (A), the Alma Mater statue (B), and a church (C) on Green Street.

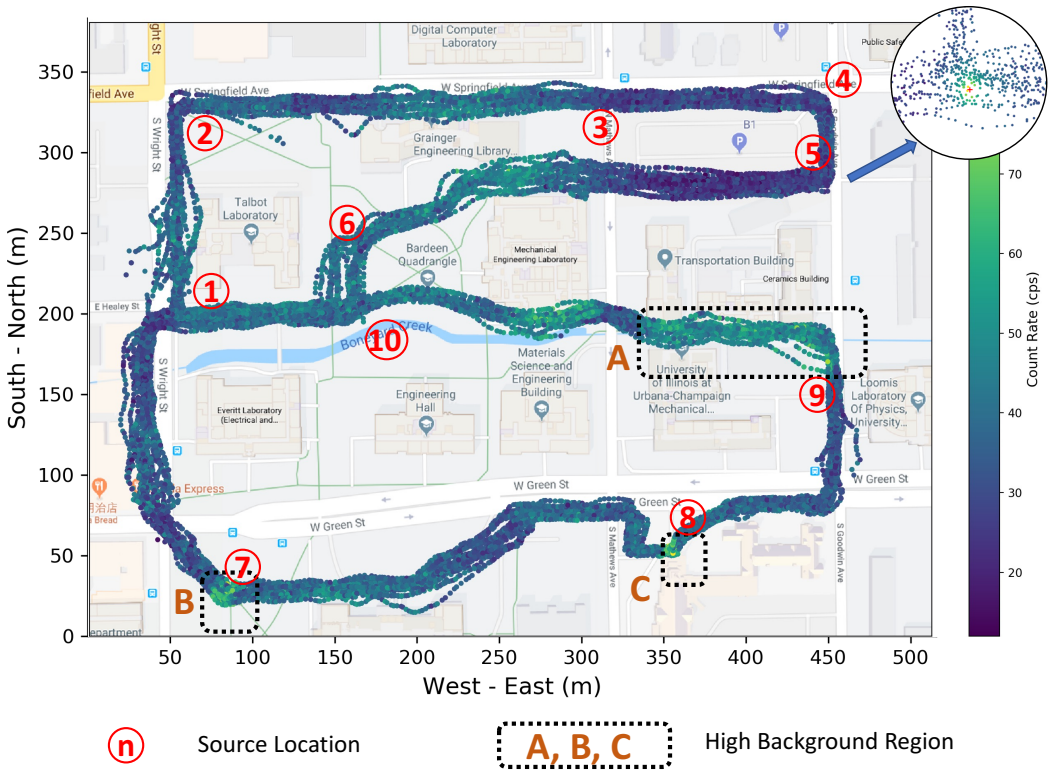


Figure F.1: Illustration of experimental area, background radiation measurements, radioactive source locations, and source injection process.

After collecting background radiation data, the synthetic radioactive sources are injected into the collected data. The radiation data collected in the morning are used to model the background radiation distribution and the radiation data collected in the afternoon are used to inject simulated radioactive sources. Ten selected locations for injecting radioactive sources are marked using red numbers from 1 to 10. After synthetic source injections, the measurements that are close to the injected radioactive sources are influenced by the source and their radiation count rates are increased. The amplitude of the increase is determined by the source intensity, shielding, and distance between radiation detector and sources. The recorded data contain the detector's id, latitude, longitude, UNIX time, gross counts, and the corresponding spectrum. The collected background radiation data and the simulated data can be accessed through Illinois Data Bank at the following address: `https://doi.org/10.13012/B2IDB-9119873_V1`. A snapshot of the collected data is shown in Fig. F.2.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	index	instrument	lat	lon	alt	error	time	duration	neutrons	counts	Spectrum					
2											s1	s2	s3	s4	s5	s6
3	0	97c-51c0-93d1-8	40.111873	-88.228772	190	20	1449849601	1002	0	35	0	0	0	0	0	0
4	1	97c-51c0-93d1-8	40.111873	-88.228772	190	20	1449849602	1000	0	51	0	0	0	0	0	0
5	2	97c-51c0-93d1-8	40.1118731	-88.228772	190	20	1449849603	1005	0	42	0	0	0	0	0	0
6	3	97c-51c0-93d1-8	40.1118731	-88.228771	190	20	1449849604	998	0	39	0	0	0	0	0	0
7	4	97c-51c0-93d1-8	40.1118731	-88.228771	190	20	1449849605	1002	0	47	0	0	0	0	0	0
8	5	97c-51c0-93d1-8	40.1118732	-88.228771	190	20	1449849606	1036	0	45	0	0	0	0	0	0
9	6	97c-51c0-93d1-8	40.1118732	-88.228771	190	20	1449849607	964	0	36	0	0	0	0	0	0
10	7	97c-51c0-93d1-8	40.1118733	-88.228771	190	20	1449849608	1001	0	44	0	0	0	0	0	0
11	8	97c-51c0-93d1-8	40.1118733	-88.228771	190	20	1449849609	1000	0	19	0	0	0	0	0	0
12	9	97c-51c0-93d1-8	40.1118733	-88.22877	190	20	1449849610	1000	0	29	0	0	0	0	0	0
13	10	97c-51c0-93d1-8	40.1118734	-88.22877	190	20	1449849611	1000	0	21	0	0	0	0	0	0
14	11	97c-51c0-93d1-8	40.1118734	-88.22877	190	20	1449849612	1003	0	32	0	0	0	0	0	0
15	12	97c-51c0-93d1-8	40.1118734	-88.22877	190	20	1449849613	998	0	22	0	0	0	0	0	0
16	13	97c-51c0-93d1-8	40.1118735	-88.22877	190	20	1449849614	1002	0	14	0	0	0	0	0	0
17	14	97c-51c0-93d1-8	40.1118735	-88.228769	190	20	1449849615	1000	0	16	0	0	0	0	0	0
18	15	97c-51c0-93d1-8	40.1118736	-88.228769	190	20	1449849616	1001	0	21	0	0	0	0	0	0
19	16	97c-51c0-93d1-8	40.1118736	-88.228769	190	20	1449849617	1002	0	19	0	0	0	0	0	0
20	17	97c-51c0-93d1-8	40.1118736	-88.228769	190	20	1449849618	999	0	22	0	0	0	0	0	0
21	18	97c-51c0-93d1-8	40.1118737	-88.228769	190	20	1449849619	1000	0	28	0	0	0	0	0	0
22	19	97c-51c0-93d1-8	40.1118737	-88.228768	190	20	1449849620	1001	0	19	0	0	0	0	0	0
23	20	97c-51c0-93d1-8	40.1118737	-88.228768	190	20	1449849621	1001	0	11	0	0	0	0	0	0
24	21	97c-51c0-93d1-8	40.1118738	-88.228768	190	20	1449849622	1002	0	24	0	0	0	0	0	0
25	22	97c-51c0-93d1-8	40.1118738	-88.228768	188	114	1449849623	1000	0	34	0	0	0	0	0	0
26	23	97c-51c0-93d1-8	40.1118739	-88.228768	188	114	1449849624	1027	0	28	0	0	0	0	0	0
27	24	97c-51c0-93d1-8	40.1118739	-88.228768	188	114	1449849625	974	0	23	0	0	0	0	0	0
28	25	97c-51c0-93d1-8	40.1118739	-88.228767	188	114	1449849626	1001	0	25	0	0	0	0	0	0
29	26	97c-51c0-93d1-8	40.111874	-88.228767	188	114	1449849627	1000	0	25	0	0	0	0	0	0
30	27	97c-51c0-93d1-8	40.111874	-88.228767	186	39	1449849628	1003	0	20	0	0	0	0	0	0

Figure F.2: Snapshot of collected experimental data.

Appendix G

Possion Kriging Error Case Analysis

As shown in Fig. 5.11, the proposed algorithm cannot correctly locate the radioactive source under certain conditions. One of the major reasons is that there exist some difference between the data (background radiation data collected in the morning) used to estimate background radiation levels and the data (background radiation data collected in the afternoon) used to inject radioactive sources. Considering the measured data around the statue, some statistical characteristics are shown in Table G.1 for the background radiation data collected in the morning and the data collected in the afternoon. Clearly, radiation data collected in the afternoon have larger mean compared with the data collected in the morning. This is one of the reason leading to the results shown in Fig. 5.11.

Table G.1: Statistical Distribution of Radiation Data from Morning and Afternoon

Measurement (*cps*)	**Min**	**25%**	**50%**	**75%**	**Max**	**Mean**	**Standard Deviation**
Morning	20.00	37.00	48.00	57.00	83.00	48.37	13.17
Afternoon	26.00	45.50	53.00	62.00	73.00	53.10	12.01

To better support this analysis, the same source injection procedure is repeated using the background radiation data collected in the morning, instead of using the data collected in the afternoon. The proposed SNR-based anomaly score is calculated as described in Chapter 5, and the results are shown in Fig. G.1. Compared with Fig. 5.11, using the data collected in the morning for source injection, the proposed algorithm can correctly locate the radioactive source. This supports the explanation behind Fig. 5.11.

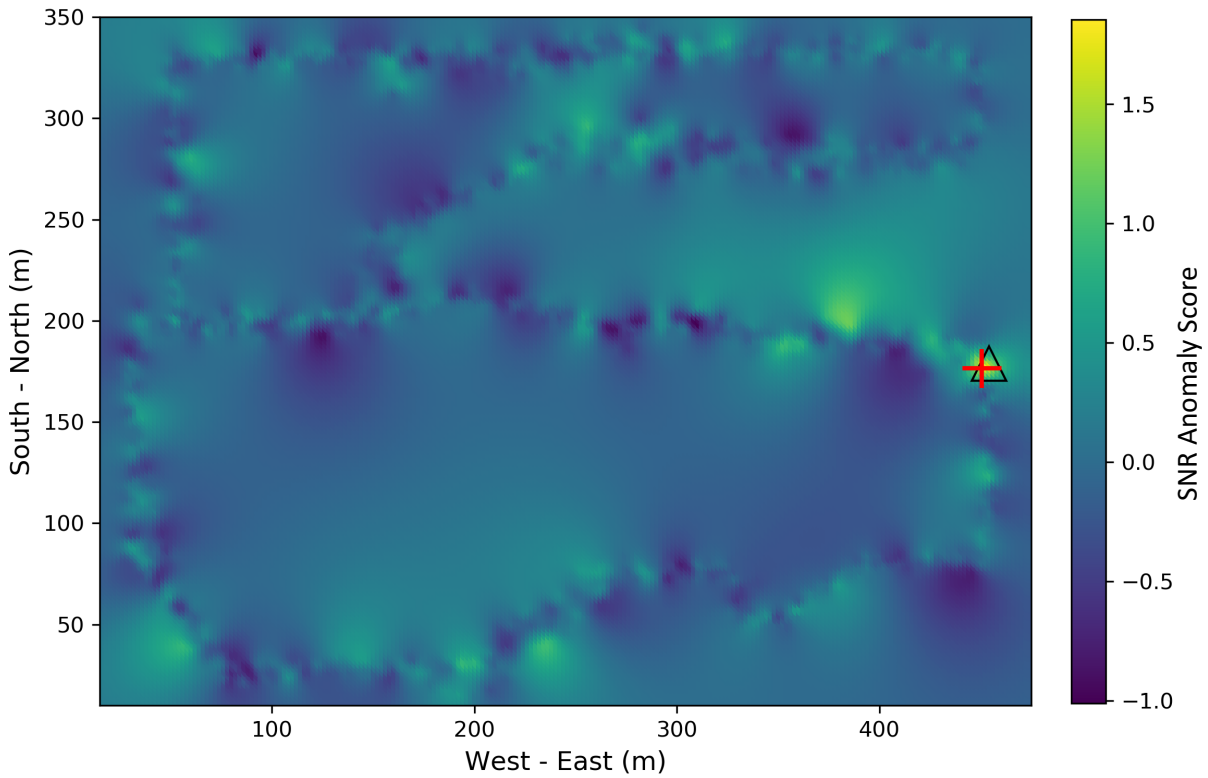


Figure G.1: A case where source injection is conducted on the data collected in the morning. The injected source is denoted by the red cross and the estimated source location is denoted by the black triangle.

Appendix H

Proof that Profile Log-Likelihood Function is Concave

Here, following the procedure shown in Reference [68], the proof is rewritten in the notations used in this work.

The profile log-likelihood function $\ell_{\boldsymbol{r}_0}(s_0, b_0)$ is denoted by:

$$\ell_{\boldsymbol{r}_0}(s_0, b_0) = \sum_{i=1}^{N} \left[m_i \cdot log(b_0 + s_i) - (b_0 + s_i) - log(m_i!) \right], \tag{H.1}$$

where s_0 is the source intensity, b_0 is the background radiation level, s_i is the average source intensity at the position of i-th detector, and m_i is the radiation count rate measurement from i-th detector.

For above equation, a radioactive source with source intensity s_0 is assumed to be placed at location $\boldsymbol{r}_0$, the i-th detector is placed at position $\boldsymbol{r}_i$, and their distance is $d_i = ||\boldsymbol{r}_i - \boldsymbol{r}_0||$. The first-order and the second-order partial derivatives of the profile log-likelihood function are give below.

The first-order partial derivatives are:

$$\frac{\partial \ell_{\boldsymbol{r}_0}}{\partial s_0} = \sum_{i=1}^{N} \left[\frac{m_i}{b_0 d_i^2 + s_0} - \frac{1}{d_i^2} \right] \tag{H.2}$$

$$\frac{\partial \ell_{\boldsymbol{r}_0}^2}{\partial b_0} = \sum_{i=1}^{N} \left[\frac{m_i d_i^2}{b_0 d_i^2 + s_0} - 1 \right]. \tag{H.3}$$

The second-order partial derivatives are:

$$\frac{\partial \ell_{\boldsymbol{r}_0}^2}{\partial {s_0}^2} = -\sum_{i=1}^{N} \frac{m_i}{(b_0 d_i^2 + s_0)^2} \tag{H.4}$$

$$\frac{\partial \ell_{\boldsymbol{r}_0}}{\partial {b_0}^2} = -\sum_{i=1}^{N} \frac{m_i d_i^4}{(b_0 d_i^2 + s_0)^2} \tag{H.5}$$

$$\frac{\partial \ell_{\boldsymbol{r}_0}^2}{\partial s_0 \partial b_0} = -\sum_{i=1}^{N} \frac{m_i d_i^2}{(b_0 d_i^2 + s_0)^2}. \tag{H.6}$$

The Hessian matrix of the profile log-likelihood function is:

$$\Delta_{\boldsymbol{\theta}}(\ell_{\boldsymbol{r}_0}) = \begin{pmatrix} \frac{\partial \ell^2_{\boldsymbol{r}_0}}{\partial {s_0}^2} & \frac{\partial \ell^2_{\boldsymbol{r}_0}}{\partial s_0 \partial b_0} \\ \\ \frac{\partial \ell^2_{\boldsymbol{r}_0}}{\partial b_0 \partial s_0} & \frac{\partial \ell^2_{\boldsymbol{r}_0}}{\partial {b_0}^2} \end{pmatrix}. \tag{H.7}$$

The determinant of Hessian matrix $\Delta_{\boldsymbol{\theta}}(\ell_{\boldsymbol{r}_0})$ is:

$$det\left(\Delta_{\boldsymbol{\theta}}(\ell_{\boldsymbol{r}_0})\right) = \left[\sum_{i=1}^{N} \frac{m_i}{(b_0 d_i^2 + s_0)^2}\right]\left[\sum_{i=1}^{N} \frac{m_i d_i^4}{(b_0 d_i^2 + s_0)^2}\right] - \left[\sum_{i=1}^{N} \frac{m_i d_i^2}{(b_0 d_i^2 + s_0)^2}\right]^2. \tag{H.8}$$

By Cauchy-Schwartz inequality, it can be shown that:

$$det\left(\Delta_{\boldsymbol{\theta}}(\ell_{\boldsymbol{r}_0})\right) \geq 0. \tag{H.9}$$

Since both $\frac{\partial \ell^2_{\boldsymbol{r}_0}}{\partial {s_0}^2}$ and $\frac{\partial \ell^2_{\boldsymbol{r}_0}}{\partial {b_0}^2}$ are negative, the profile log-likelihood function $\ell_{\boldsymbol{r}_0}(s_0, b_0)$ is concave in s_0 and b_0.

Appendix I

Maximum Likelihood Estimation Optimization Techniques

The log-likelihood function is shown in Eq. I.1.

$$\begin{aligned}\ell(\boldsymbol{\theta}) = \ell(s_0, \boldsymbol{r}_0, b_0) &= \sum_{i=1}^{N} \left[m_i \cdot log(b_0 + s_i) - (b_0 + s_i) - log(m_i!) \right] \\ &= \sum_{i=1}^{N} \left[m_i \cdot log(b_0 + \frac{s_0}{||\boldsymbol{r}_i - \boldsymbol{r}_0||^2}) - (b_0 + \frac{s_0}{||\boldsymbol{r}_i - \boldsymbol{r}_0||^2}) - log(m_i!) \right].\end{aligned} \tag{I.1}$$

To maximize the profile log-likelihood function $\ell_{\boldsymbol{r}_0}(\boldsymbol{\theta})$, gradient ascent method or Newton-Raphson method can be used.

I.1 Gradient Ascent Method

The iterative scheme for gradient ascent method is:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \eta \nabla_{\boldsymbol{\theta}^{(k)}} \left(\ell_{\boldsymbol{r}_0}(\boldsymbol{\theta}^{(k)}) \right), \tag{I.2}$$

where k is the number of calculation iterations, η is learning rate, and $\nabla_{\boldsymbol{\theta}^{(k)}}$ is the gradient.

So, if the background radiation level b_i is assumed to be known in advance, then the expression for s_0 is:

$$s_0^{(k+1)} = s_0^{(k)} + \eta \sum_{i=1}^{N} \left[\frac{m_i}{b_i d_i^2 + s_0^{(k)}} - \frac{1}{d_i^2} \right]. \tag{I.3}$$

In Eq. I.3, background radiation level b_i is assumed to be known. When b_i is unknown and the study area is not large, b_i can be assumed to be same across the region, denoted by b_0. Then, the expression for $s_0^{(k+1)}$ and $b_0^{(k+1)}$ become:

$$\begin{pmatrix} s_0^{(k+1)} \\ \\ b_0^{(k+1)} \end{pmatrix} = \begin{pmatrix} s_0^{(k)} \\ \\ b_0^{(k)} \end{pmatrix} + \eta \begin{pmatrix} \sum_{i=1}^{N} \left[\frac{m_i}{b_0^{(k)} d_i^2 + s_0^{(k)}} - \frac{1}{d_i^2} \right] \\ \\ \sum_{i=1}^{N} \left[\frac{m_i d_i^2}{b_0^{(k)} d_i^2 + s_0^{(k)}} - 1 \right] \end{pmatrix}. \tag{I.4}$$

I.2 Newton-Raphson Method

The iterative scheme for the Newton-Raphson algorithm is:

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \Delta_{\boldsymbol{\theta}^{(k)}}^{-1} \left(\ell_{\boldsymbol{r}_0} \left(\boldsymbol{\theta}^{(k)} \right) \right) \nabla_{\boldsymbol{\theta}^{(k)}} \left(\ell_{\boldsymbol{r}_0} \left(\boldsymbol{\theta}^{(k)} \right) \right), \tag{I.5}$$

where $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{r}_0 \\ \\ s_0 \end{pmatrix}$ is the vector to be estimated, $\ell_{\boldsymbol{r}_0}(\boldsymbol{\theta}^{(k)})$ is the profile log-likelihood function, $\nabla_{\boldsymbol{\theta}^{(k)}}$ is the gradient, and $\Delta_{\boldsymbol{\theta}^{(k)}}$ is the Hessian matrix [68].

Given some initial guess $\boldsymbol{\theta}^{(0)}$, the MLE of $\boldsymbol{\theta}$ can be iteratively found by using Eq. I.5. More precisely, based on Eq. 6.5, terms appearing in Eq. I.5 can be written as:

$$\nabla_{\boldsymbol{\theta}}(\ell_{\boldsymbol{r}_0}) = \begin{pmatrix} \frac{\partial \ell_{\boldsymbol{r}_0}}{\partial s_0} \\ \\ \frac{\partial \ell_{\boldsymbol{r}_0}}{\partial b_0} \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{N} \left[\frac{m_i}{b_0 d_i^2 + s_0} - \frac{1}{d_i^2} \right] \\ \\ \sum\limits_{i=1}^{N} \left[\frac{m_i d_i^2}{b_0 d_i^2 + s_0} - 1 \right] \end{pmatrix} \tag{I.6}$$

$$\Delta_{\boldsymbol{\theta}}(\ell_{\boldsymbol{r}_0}) = \begin{pmatrix} \frac{\partial \ell_{\boldsymbol{r}_0}^2}{\partial {s_0}^2} & \frac{\partial \ell_{\boldsymbol{r}_0}^2}{\partial s_0 \partial b_0} \\ \\ \frac{\partial \ell_{\boldsymbol{r}_0}^2}{\partial b_0 \partial s_0} & \frac{\partial \ell_{\boldsymbol{r}_0}^2}{\partial {b_0}^2} \end{pmatrix} = \begin{pmatrix} -\sum\limits_{i=1}^{N} \frac{m_i}{(b_0 d_i^2 + s_0)^2} & -\sum\limits_{i=1}^{N} \frac{m_i d_i^2}{(b_0 d_i^2 + s_0)^2} \\ \\ -\sum\limits_{i=1}^{N} \frac{m_i d_i^2}{(b_0 d_i^2 + s_0)^2}] & -\sum\limits_{i=1}^{N} \frac{m_i d_i^4}{(b_0 d_i^2 + s_0)^2} \end{pmatrix} \tag{I.7}$$

$$\Delta_{\boldsymbol{\theta}}^{-1}(\ell_{\boldsymbol{r}_0}) = \frac{1}{|\Delta_{\boldsymbol{\theta}(\ell_{\boldsymbol{r}_0})}|} \begin{pmatrix} -\sum_{i=1}^{N} \frac{m_i d_i^4}{(b_0 d_i^2 + s_0)^2} & \sum_{i=1}^{N} \frac{m_i d_i^2}{(b_0 d_i^2 + s_0)^2} \\ \sum_{i=1}^{N} \frac{m_i d_i^2}{(b_0 d_i^2 + s_0)^2} & -\sum_{i=1}^{N} \frac{m_i}{(b_0 d_i^2 + s_0)^2} \end{pmatrix}. \tag{I.8}$$

During practical implementation, it is found that with appropriate initialization, for example, setting s_0 to be 0 and b_0 to be 38 *cps* for D3S detectors, Newton-Raphson algorithm converges within 50 iterations, while gradient ascent method need more than 1,000 iterations to converge if the learning rate is set to be 0.01. Thus, in this work, the Newton-Raphson algorithm is implemented for optimization.

Appendix J

Preliminary Analysis for Mobile Radioactive Sources

Chapter 6 describes the application of KDE with MLE and grid search with MLE for cases in which radioactive sources are stationary. When the radioactive sources are stationary, since there will be continuous measurements around the radioactive sources, the calculation can be continued as new data are collected, which leads to relatively accurate results. However, in other cases, radioactive sources might be carried by people or vehicles, which makes the source localization task much more difficult. Compared with stationary radioactive sources, for mobile radioactive sources, one of the most important difference is that there is no guarantee that the continuous measurements from the region surrounding the radioactive sources will be available, especially when there are limited number of detectors. It is possible that all the detectors at some point of time are far away from the radioactive sources, which makes the radioactive sources statistically impossible to be detected based on the measurements at that time.

For a mobile source, since the radioactive source keeps moving, the approach which uses all the available measurements around the source does not work. Instead, a small time window can be selected and the measurements within that time window should be used for calculation. The smallest time window is 1, in which case only the measurements at the moment are used to locate the radioactive sources. When the chosen time window is larger than 1, some previous measurements will also be used. Theoretically, if the time window size is too small, the number of measurements used for the calculation will also be small, which restricts the performance of the proposed algorithm. On the other hand, if the time window size is too large, the mobile source might have moved a large distance during the chosen period, which will also restrict the performance of the mobile sensor network.

To study the performance of proposed algorithms for mobile sources, simulations are conducted with multiple mobile detectors and single mobile source in the study area in each simulation. The experimental setting is the same as described in section 3.3. 50 detectors move along the street with constant speed for each detector. The radioactive source is also moving along the road. Its speed is fixed at 1.4 m/s. Similar to mobile detectors, radioactive source have the same probability to move along the original direction, turn left, or turn right at the intersections. And when it reaches the boundary, it turns back and the simulation

continues. The source intensities are chosen to be 100 μCi, 300 μCi, 500 μCi, 1,000 μCi, and 2,000 μCi. The box plots[1] of the source localization error for 500 μCi sources are shown in Fig. J.1. The horizontal axis corresponds to the number of detectors, vertical axis corresponds to source localization errors, and different time windows (1 through 10) are denoted using different colors. Number of detectors in the simulation are: 40, 60, 80, $\cdots$, 160.

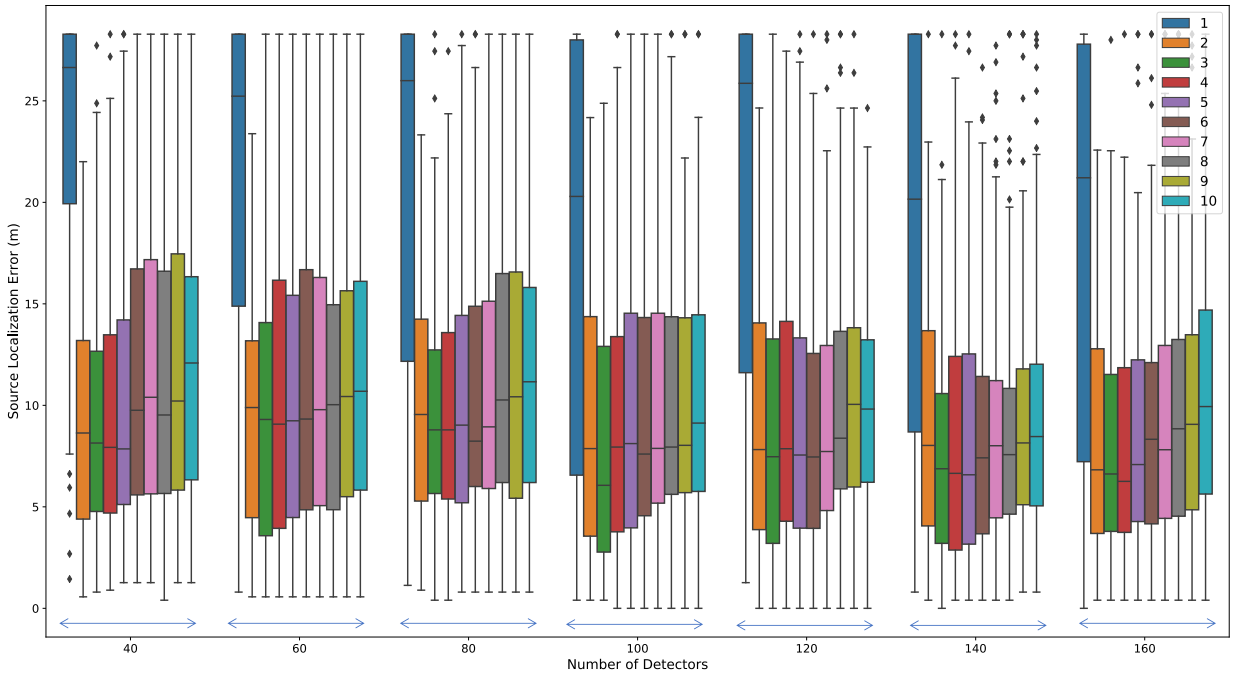


Figure J.1: Box plot for source localization errors with 500 μCi source. Time window varying from 1 second to 10 seconds is denoted using different colors. Number of detectors are varied from 40 to 160.

In Fig. J.1, the time window size varies from 1 second to 10 seconds, and the source localization error varies as a function of number of detectors and the time window size. When the time window is only 1 second, the source localization error is very large. Error drops dramatically from time window of 2 seconds, but then remains fairly uniform (and in fact, increasing slightly in some cases with larger time windows). Considering the distribution of the source localization error and for the conditions studied here, time windows of 3, 4, or 5 seconds seem to be optimal choices.

Impact of detector quality and source intensity on source localization error is shown in Fig. J.2. Setting the time window to be 4 seconds, the source localization error with 60 detectors are shown in Fig J.2 for

[1]Box plot is a method to visualize the distribution of numerical data through quartiles. In the middle, the median is shown using a horizontal line. And the box shows area between 25% percentile (Q1) and 75% percentile (Q3). The IQR is defined as the interquartile range between 25% percentile and 75% percentile. The lowest horizontal line shows the value corresponding to Q1 - 1.5 $\times$ IQR and the highest horizontal line shows the value corresponding to Q3 + 1.5 $\times$ IQR. All the measurements that is beyond this range are defined as outlier, and are shown using dots.

different source intensities. Results for D3S detector and Ortec-905 detector are denoted using different colors. Not surprisingly, with strong radioactive sources, the median source localization error decreases, especially for D3S detectors. For Ortec 905 detectors, this improvement is small since for 100 μCi radioactive source, the source localization error is already small. Using better detector, for example, Ortec versus D3S, also improves the source localization results, especially for the cases with weak radioactive sources. For example, for 100 μCi radioactive source, the median source localization error is around 7 meters for Ortec 905 detectors, while it is around 10 meters for D3S detectors. But in all these cases, the performance of mobile sensor networks is much worse compared with stationary radioactive sources.

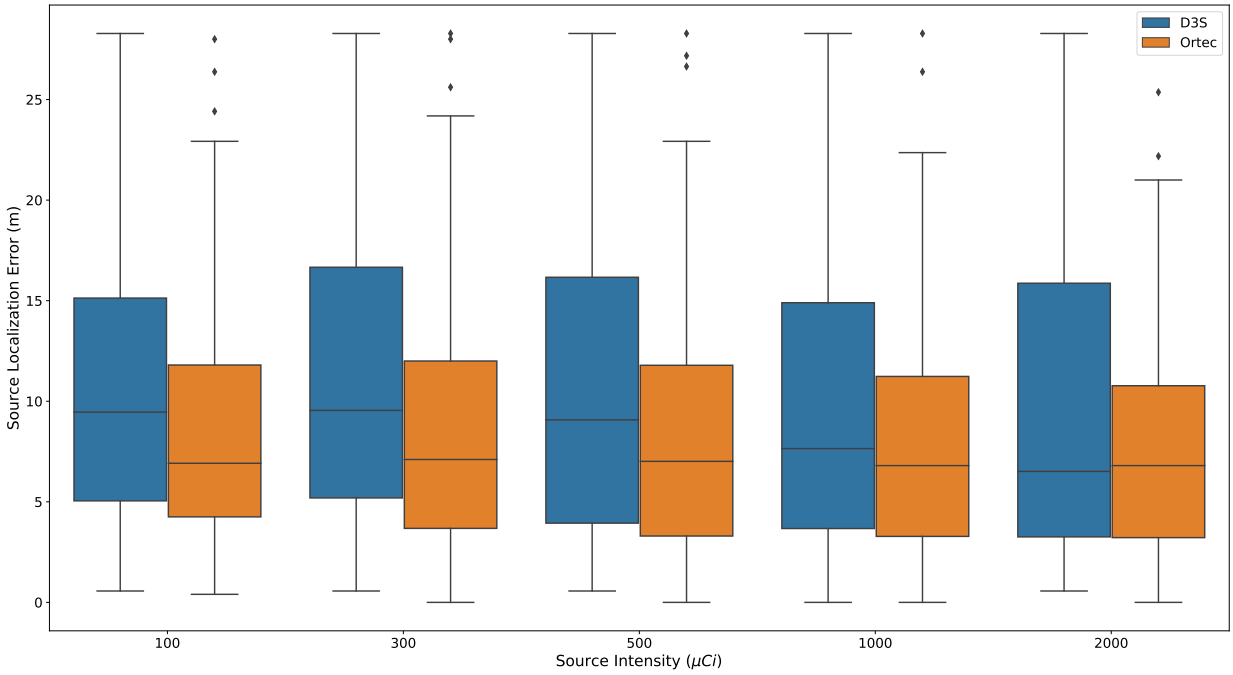


Figure J.2: Box plot for source localization errors with 60 detectors.

As discussed in this section, when radioactive sources are moving, accurately detecting or locating the radioactive sources becomes much more difficult. When there is only limited number of detectors, the probability that a detector can reach close to the mobile sources is small. Even after appropriate time window is selected, the performance of the mobile sensor network is still much worse compared with stationary radioactive sources. Further studies are needed to address this problem.

Appendix K

Connection between Grid Search with Maximum Likelihood Estimation and Bayesian Estimation

The source localization problem from the perspective of parameter estimation is discussed in Chapter 6. Parameter estimation is essentially a process of maximizing the likelihood function. KDE with MLE and grid search with MLE are two approaches to solve or approximately solve the maximization problem. This is a process that is purely based on the collected measurements. However, in real cases, different kinds of prior information can also be included to optimize the source localization process.

From the perspective of hypothesis testing [8], for each position $\boldsymbol{r}_0$ in the study area, the null hypothesis H_0 and alternative hypothesis H_a are:

H_0: There is no radioactive source at position $\boldsymbol{r}_0$.

H_a: There is radioactive source with certain strength at position $\boldsymbol{r}_0$.

Given the measurements $m_1, m_2, \cdots, m_N$ used for calculation, the goal is to calculate $P(H_0|m_1, m_2, \cdots, m_N)$ and $P(H_a|m_1, m_2, \cdots, m_N)$ such that the ratio can be used to measure the likelihood that there are radioactive sources at position $\boldsymbol{r}_0$. In other words, the position with highest value of $\frac{P(H_a|m_1,m_2,\cdots,m_N)}{P(H_0|m_1,m_2,\cdots,m_N)}$ is assumed to be the most likely source location.

Based on Bayes rule, the ratio can be derived as:

$$\frac{P(H_a|m_1, m_2, \cdots, m_N)}{P(H_0|m_1, m_2, \cdots, m_N)} = \frac{P(H_a) \cdot P(m_1, m_2, \cdots, m_N|H_a)}{P(H_0) \cdot P(m_1, m_2, \cdots, m_N|H_0)} = \frac{P(H_a)}{P(H_0)} \cdot \frac{P(m_1, m_2, \cdots, m_N|H_a)}{P(m_1, m_2, \cdots, m_N|H_0)}. \quad \text{(K.1)}$$

Since $P(m_1, m_2, \cdots, m_N|H_0)$ can be seen as constant given $m_1, m_2, \cdots, m_N$, and $P(m_1, m_2, \cdots, m_N|H_a)$ is the same as discussed in Section 6.2.1, if the effect of prior information $\frac{P(H_a)}{P(H_0)}$ is ignored, this method is equivalent to the grid search with MLE approach discussed previously. However, the prior information can be used to improve the search algorithms.

Considering $P(H_a)$, that there is radioactive source with certain strength at position $\boldsymbol{r}_0$, $P(H_a)$ is likely to be higher at crowded places and lower at other areas. Through incorporating the prior information $\frac{P(H_a)}{P(H_0)}$,

the distribution of $\frac{P(H_a|m_1,m_2,\cdots,m_N)}{P(H_0|m_1,m_2,\cdots,m_N)}$ could be different from the case in which prior information is ignored. In this way, when there may exist multiple potential radioactive sources, although all the potential regions should be examined, with limited amount of resources, the posterior ratio distribution provides the priority of searching and will help reasonably allocate the resources.

Appendix L

Sample Code for Autoencoder Implemented in TensorFlow

```
import random
import pickle
import numpy as np
import tensorflow as tf
import keras
from keras.models import Sequential, Model
from keras.layers import Dense, Dropout, Flatten
from keras.layers import Conv1D, MaxPooling1D

def autoencoder(layers=(512, 256, 128, 256, 512), name='Autoencoder'):
    """ define fully-connected Autoencoder structure """
    model = Sequential(name=name)
    model.add(Dense(input_shape=(layers[0],), units=layers[1], activation='relu', name='
    dense-1'))
    for i in range(2, len(layers)):
        if i == len(layers) - 1:
            model.add(Dense(units=layers[i], activation='relu', name='output'))
        else:
            model.add(Dense(units=layers[i], activation='relu', name='dense-' + str(i)))

    return model


# To reproduce the result, refer to https://keras.io/getting-started/faq/
seed = 42
np.random.seed(seed)
random.seed(seed)
session_conf = tf.ConfigProto(intra_op_parallelism_threads=1, inter_op_parallelism_threads
    =1)
tf.set_random_seed(seed)
sess = tf.Session(graph=tf.get_default_graph(), config=session_conf)
```

```
K.set_session(sess)

# Define autoencoder
layers=(1024, 512, 256, 512, 1024)
model = autoencoder(layers=layers, name='Autoencoder')

# In the following code, assume the training data, validation data are denoted by train, val
# Create optimizer
opt = keras.optimizers.Adam()
model.compile(optimizer=opt, loss='mean_absolute_error')

# Create call backs
checkpoint = ModelCheckpoint(filepath='./model/MAE_autoencoder-' + str(layers) + '-ckpt.h5', verbose=0, monitor='val_loss', save_best_only=True)

# Fit the model
history = model.fit(x=train, y=train, batch_size=512, epochs=500, verbose=0, callbacks=[checkpoint], validation_data=(val, val))

# Save model and history
pickle.dump(history.history, open('./model/MAE_autoencoder-' + str(layers) + '-history.p', 'wb'))
```

Listing L.1: Demo of Autoencoder Implementation

Appendix M

K-Nearest Neighbors Query Time Analysis

As discussed in section 7.4, proximity-based method are usually called "lazy" method. There is generally no procedure to fit or train a model first and make predictions later. Instead, for each unknown measurement, the proximity-based model needs to calculate the distance to all measurements in the dataset. For the dataset with N instances and D dimensions, in the the naive implementation, the query time grows as $O(N \cdot D)$, which is very expensive for cases with large N or D.

To speed up the computation for proximity-based methods, K-D Tree [103] and Ball Tree [104] can be implemented. To compare the influence of different data structures on KNN method, 62,500 spectra are used as the training dataset and the query time is calculated for different number of nearest neighbors k. The result is shown in Fig. M.1.

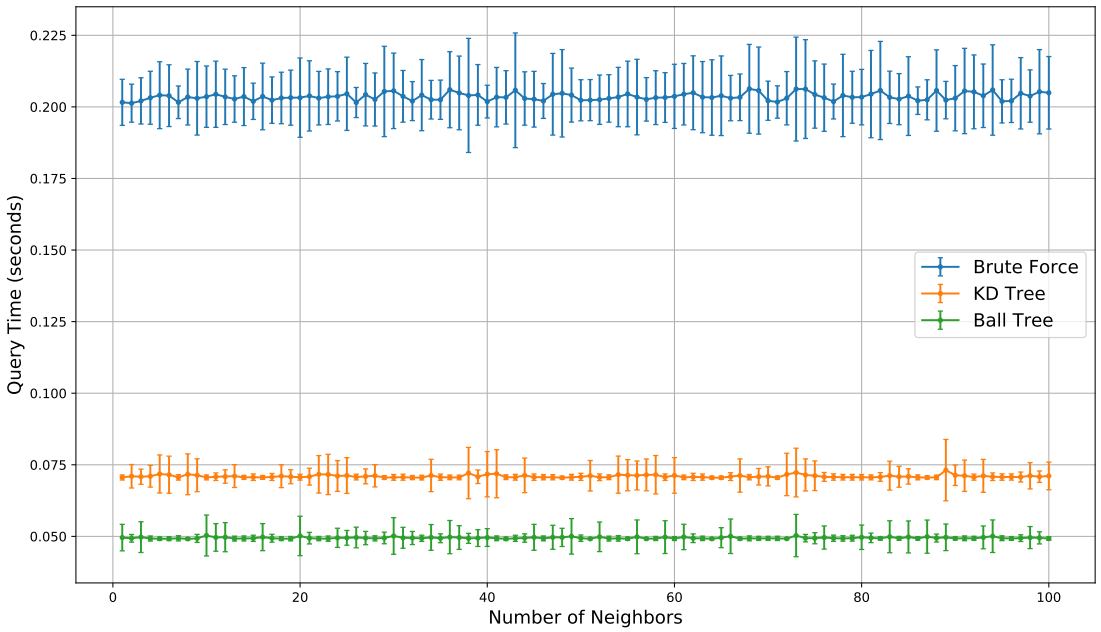


Figure M.1: Comparison of query time for KNN with different number of neighbors.

From Fig. M.1, the number of nearest neighbors to query has little influence on the total query time,

but the data structure has dramatic influence on the query time. K-D Tree and Ball Tree can dramatically speed up the query process.

In addition to data structure, another influencing factor is the sample size of the training set. Figure M.2 shows the query time for different sample size N. For all three cases, the average query time increases roughly linearly as the sample size increases. But using K-D Tree and Ball Tree can help reduce the query time, especially for large amount of training data.

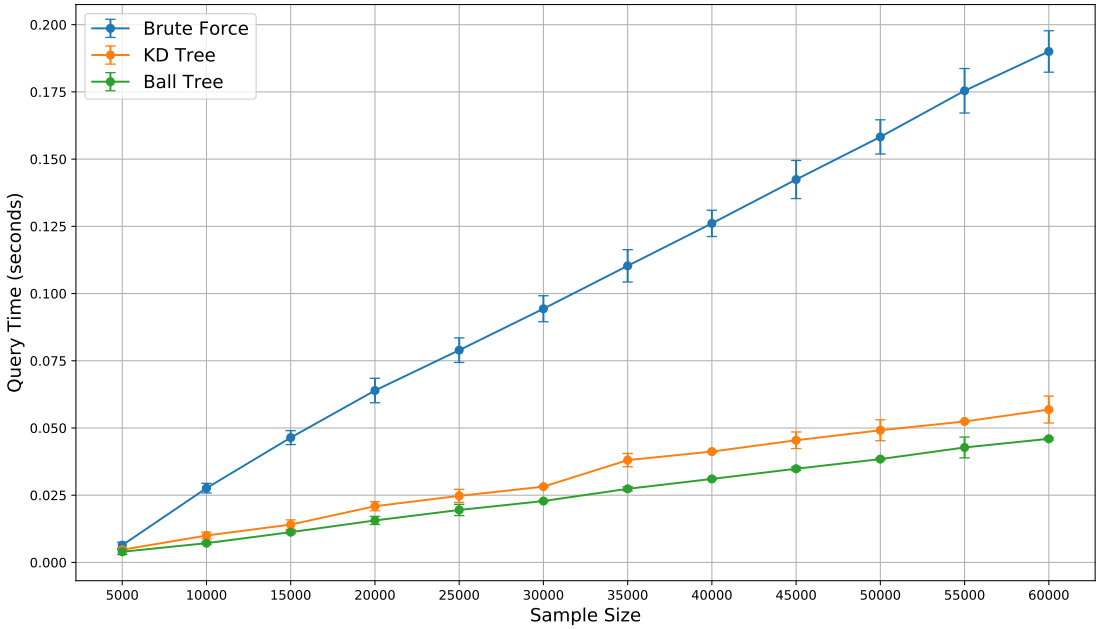


Figure M.2: Comparison of query time for KNN with different sample size.

Appendix N

Application of Anomaly Detection

As discussed in Chapter 1, anomaly detection, source localization, and isotope identification are three major problems involved in nuclear radiation detection. These three parts are interconnected. For example, source localization is usually based on the results of anomaly detection. Through anomaly detection, the search regions are restricted to relatively smaller areas, which can reduce the computational resources needed and reduce the false positive rates.

Anomaly detection, especially the spectrum-based anomaly detection, is also important for isotope identification. Through anomaly detection, the spectra that contain no information about radioactive sources are removed, which not only reduces the computational resources, but also eliminates the influence of noisy data. In addition, reconstruction-based anomaly detection techniques, for example, PCA or autoencoder, can make isotope identification task easier.

Considering three isotopes, ^{60}Co, ^{137}Cs, and ^{54}Mn, using PCA, the dimension of the spectra can be reduced from 512 to 2, and three different spectra can be visualized in two dimensions as shown in Fig. N.1.

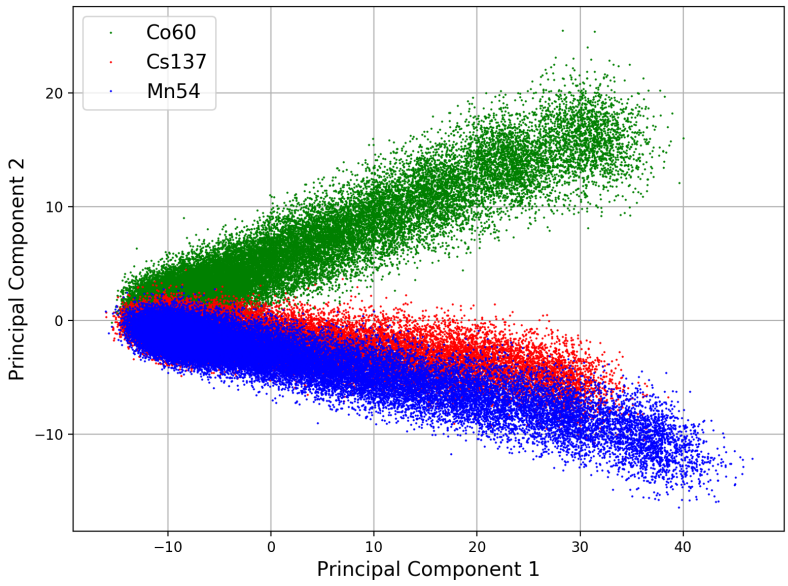
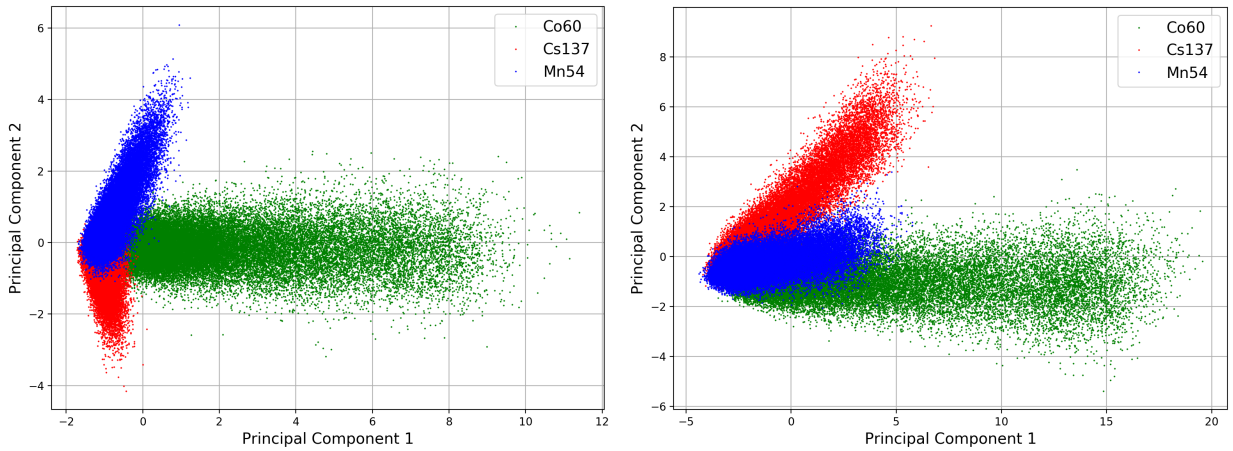


Figure N.1: Visualization of sample ^{60}Co, ^{137}Cs, and ^{54}Mn spectra PCA projection. Data are simulated using GADRAS as described in Chapter 7.

As shown in Fig. N.1, the spectra of ^{137}Cs and ^{54}Mn are quite similar to each other. When source intensities are not strong, background radiation spectra also blur the boundary between different isotopes. To make it easier for different isotopes being separated with each other, PCA and autoencoder described in section 7.2 can be applied to estimate background radiation spectra through reconstructing the original radiation spectra. After estimating the background spectra, the difference between original spectra and reconstructed spectra are assumed to mainly contain the information from isotopes instead of background radiation. Through PCA, the difference spectra can be visualized again using the projection on the first two principal component directions. The results are shown in Fig. N.2.



(a) Projection after PCA.

(b) Projection after Autoencoder.

Figure N.2: Visualization of sample ^{60}Co, ^{137}Cs, and ^{54}Mn spectra PCA projection after dimensionality reduction and reconstruction.

As shown clearly, after dimensionality reduction and reconstruction using PCA and autoencoder, the three isotopes are better separated compared with the raw spectra in Fig. N.1. The isotope identification methods, for example, the methods proposed by Kamuda et al. [35], can be directly applied to the difference spectra instead of the original radiation spectra.

References

[1] S. M. Brennan, A. M. Mielke, and D. C. Torney. Radioactive source detection by sensor networks. *IEEE Transactions on Nuclear Science*, 52(3):813–819, 2005.

[2] S. M. Brennan, A. M. Mielke, D. C. Torney, and A. B. MacCabe. Radiation detection with distributed sensor networks. *Computer*, 37(8):57–59, 2004.

[3] R. J. Nemzek, J. S. Dreicer, D. C. Torney, and T. T. Warnock. Distributed sensor networks for detection of mobile radioactive sources. *IEEE Transactions on Nuclear Science*, 51(4):1693–1700, 2004.

[4] J. Zhao. Radiation source detection from mobile sensor networks using principal component analysis. Master's thesis, University of Illinois at Urbana-Champaign, 2016.

[5] M. Chandy, C. Pilotto, and R. McLean. Networked sensing systems for detecting people carrying radioactive material. In *Networked Sensing Systems, 2008. INSS 2008. 5th International Conference on*, pages 148–155. IEEE, 2008.

[6] K. M. Chandy, J. Bunn, and A. Liu. Models and algorithms for radiation detection. In *Modeling and Simulation Workshop for Homeland Security*, pages 1–6. Citeseer, 2010.

[7] C. J. Sullivan. Nuclear forensics driven by geographic information systems and big data analytics. In *Proceedings of the Information Analysis Technologies, Techniques and Methods for Safeguards, Nonproliferation and Arms Control Verification Workshop, Portland, OR*, pages 12–14, 2014.

[8] P. Tandon. *Bayesian aggregation of evidence for detection and characterization of patterns in multiple noisy observations.* PhD thesis, Carnegie Mellon University, 2015.

[9] P. Tandon, P. Huggins, R. Maclachlan, A. Dubrawski, K. Nelson, and S. Labov. Detection of radioactive sources in urban scenes using Bayesian aggregation of data from mobile spectrometers. *Information Systems*, 57:195–206, 2016.

[10] A. H. Liu, J. J. Bunn, and K. M. Chandy. An analysis of data fusion for radiation detection and localization. In *Information Fusion (FUSION), 2010 13th Conference on*, pages 1–8. IEEE, 2010.

[11] A. H. Liu. Simulation and implementation of distributed sensor network for radiation detection. Master's thesis, California Institute of Technology, 2010.

[12] Y. T. Yang, B. Fishbain, D. S. Hochbaum, E. B. Norman, and E. Swanberg. The supervised normalized cut method for detecting, classifying, and identifying special nuclear materials. *INFORMS Journal on Computing*, 26(1):45–58, 2013.

[13] W. Bertozzi and R. J. Ledoux. Nuclear resonance fluorescence imaging in non-intrusive cargo inspection. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 241(1):820–825, 2005.

[14] E. B. Norman, S. G. Prussin, R. Larimer, H. Shugart, E. Browne, A. R. Smith, R. J. McDonald, H. Nitsche, P. Gupta, M. I. Frank, et al. Signatures of fissile materials: high-energy γ rays following fission. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 521(2):608–610, 2004.

[15] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

[16] C. C. Aggarwal. *Outlier analysis*. Springer Publishing Company, Incorporated, 2013.

[17] D. M. Hawkins. *Identification of outliers*, volume 11. Springer, 1980.

[18] A. H. Liu, J. J. Bunn, and K. M. Chandy. Sensor networks for the detection and tracking of radiation and other threats in cities. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, pages 1–12. IEEE, 2011.

[19] M. R. Morelande and B. Ristic. Radiological source detection and localisation using Bayesian techniques. *IEEE Transactions on Signal Processing*, 57(11):4220–4231, 2009.

[20] K. D. Jarman, L. E. Smith, D. K. Carlson, and D. N. Anderson. Sequential probability ratio test for long-term radiation monitoring. In *Nuclear Science Symposium Conference Record, 2003 IEEE*, volume 2, pages 1458–1462. IEEE, 2003.

[21] J.-C. Chin, N. S. V. Rao, D. K. Y. Yau, M. Shankar, Y. Yang, J. C. Hou, S. Srivathsan, and S. Iyengar. Identification of low-level point radioactive sources using a sensor network. *ACM Transactions on Sensor Networks (TOSN)*, 7(3):21, 2010.

[22] N. S. V. Rao, J.-C. Chin, D. K. Y. Yau, and C. Y. T. Ma. Localization leads to improved distributed detection under non-smooth distributions. In *Information Fusion (FUSION), 2010 13th Conference on*, pages 1–8. IEEE, 2010.

[23] J. M. Kirkpatrick and B. M. Young. Poisson statistical methods for the analysis of low-count gamma spectra. *IEEE Transactions on Nuclear Science*, 56(3):1278–1282, 2009.

[24] R. Fischer, K. M. Hanson, V. Dose, and W. von Der Linden. Background estimation in experimental spectra. *Physical Review E*, 61(2):1152, 2000.

[25] M. Morháč, J. Kliman, V. Matoušek, M. Veselskỳ, and I. Turzo. Background elimination methods for multidimensional coincidence γ-ray spectra. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 401(1):113–132, 1997.

[26] M. Morháč. An algorithm for determination of peak regions and baseline elimination in spectroscopic data. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 600(2):478–487, 2009.

[27] M. H. Zhu, L. G. Liu, Y. S. Cheng, T. K. Dong, Z. You, and A. A. Xu. Iterative estimation of the background in noisy spectroscopic data. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 602(2):597–599, 2009.

[28] M. Alamaniotis, J. Mattingly, and L. H. Tsoukalas. Kernel-based machine learning for background estimation of NaI low-count gamma-ray spectra. *IEEE Transactions on Nuclear Science*, 60(3):2209–2221, 2013.

[29] C. Fraschini and F. Chaillan. Background spectrum estimation via robust Kalman filtering. In *New Trends for Environmental Monitoring Using Passive Systems, 2008*, pages 1–6. IEEE, 2008.

[30] L. V. East, R. L. Phillips, and A. R. Strong. A fresh approach to NaI scintillation detector spectrum analysis. *Nuclear Instruments and Methods in Physics Research*, 193(1-2):147–155, 1982.

[31] M. S. Mitra and P. K. Sarkar. Monte Carlo simulations to estimate the background spectrum in a shielded NaI(Tl) γ-spectrometric system. *Applied radiation and isotopes*, 63(4):415–422, 2005.

[32] R. C. Runkle, M. F. Tardiff, K. K. Anderson, D. K. Carlson, and L. E. Smith. Analysis of spectroscopic radiation portal monitor data using principal components analysis. *IEEE Transactions on Nuclear Science*, 53(3):1418–1423, 2006.

[33] D. Boardman, M. Reinhard, and A. Flynn. Principal component analysis of gamma-ray spectra for radiation portal monitors. *IEEE Transactions on Nuclear Science*, 59(1):154–160, 2012.

[34] M. Kamuda, J. B. Stinnett, and C. J. Sullivan. Automated isotope identification algorithm using artificial neural networks. *IEEE Transactions on Nuclear Science*, 64(7):1858–1864, 2017.

[35] M. Kamuda, J. Zhao, and K. Huff. A comparison of machine learning methods for automated gamma-ray spectroscopy. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2018.

[36] J.-C. Chin, D. K Y. Yau, N. S. V. Rao, Y. Yang, C. Y. T. Ma, and M. Shankar. Accurate localization of low-level radioactive source under noise and measurement errors. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, pages 183–196. ACM, 2008.

[37] M. Morelande, B. Ristic, and A. Gunatilaka. Detection and parameter estimation of multiple radioactive sources. In *Information Fusion, 2007 10th International Conference on*, pages 1–7. IEEE, 2007.

[38] Z. Liu. Mobile radiation sensor networks for source detection in a fluctuating background using geotagged count rate data. Master's thesis, University of Illinois at Urbana-Champaign, 2016.

[39] Z. Liu, S. Abbaszadeh, and C. J. Sullivan. Spatial-temporal modeling of background radiation using mobile sensor networks. *PloS one*, 13(10):e0205092, 2018.

[40] J. Zhao and C. J. Sullivan. Detection and parameter estimation of radioactive sources with mobile sensor networks. *Radiation Physics and Chemistry*, 155:265–270, 2019.

[41] U. Bhattacharyya and C. Baum. Estimating the location of a nuclear source in a three-dimensional environment using a two-stage adaptive algorithm. In *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 8–14, Jan 2018.

[42] S. R. Dalal and B. Han. Detection of radioactive material entering national ports: A Bayesian approach to radiation portal data. *The Annals of Applied Statistics*, pages 1256–1271, 2010.

[43] K. D. Jarman, E. A. Miller, R. S. Wittman, and C. J. Gesh. Bayesian radiation source localization. *Nuclear technology*, 175(1):326–334, 2011.

[44] M. R. Morelande, C. M. Kreucher, and K. Kastella. A Bayesian approach to multiple target detection and tracking. *IEEE Transactions on Signal Processing*, 55(5):1589–1604, 2007.

[45] S. S. Ram and V. V. Veeravalli. Localization and intensity tracking of diffusing point sources using sensor networks. In *Global Telecommunications Conference, 2007. GLOBECOM'07. IEEE*, pages 3107–3111. IEEE, 2007.

[46] D. S. Hochbaum. The multi-sensor nuclear threat detection problem. In *Operations Research and Cyber-Infrastructure*, pages 389–399. Springer, 2009.

[47] D. S. Hochbaum and B. Fishbain. Nuclear threat detection with mobile distributed sensor networks. *Annals of Operations Research*, 187(1):45–63, 2011.

[48] N. S. V. Rao, M. Shankar, J.-C. Chin, D. K. Y. Yau, C. Y. T. Ma, Y. Yang, J. C. Hou, X. Xu, and S. Sahni. Localization under random measurements with application to radiation sources. In *Information Fusion, 2008 11th International Conference on*, pages 1–8. IEEE, 2008.

[49] J. M. Hite, J. K. Mattingly, K. L. Schmidt, R. Ştefănescu, and R. Smith. Bayesian metropolis methods applied to sensor networks for radiation source localization. In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2016 IEEE International Conference on*, pages 389–393. IEEE, 2016.

[50] J. W. Howse, L. O. Ticknor, and K. R. Muske. Least squares estimation techniques for position tracking of radioactive sources. *Automatica*, 37(11):1727–1737, 2001.

[51] B. Ristic, A. Gunatilaka, and M. Rutten. An information gain driven search for a radioactive point source. In *2007 10th International Conference on Information Fusion*, pages 1–8. IEEE, 2007.

[52] B. Ristic and A. Gunatilaka. Information driven localisation of a radiological point source. *Information fusion*, 9(2):317–326, 2008.

[53] B. Ristic, M. Morelande, and A. Gunatilaka. Information driven search for point sources of gamma radiation. *Signal Processing*, 90(4):1225–1239, 2010.

[54] H. Lin et al. Search strategy of a mobile robot for radiation sources in an unknown environment. In *2014 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*, pages 56–60. IEEE, 2014.

[55] H. Ardiny, S. Witwicki, and F. Mondada. Autonomous exploration for radioactive hotspots localization taking account of sensor limitations. *Sensors*, 19(2):292, 2019.

[56] Z. Liu and S. Abbaszadeh. Double Q-Learning for radiation source detection. *Sensors*, 19(4):960, 2019.

[57] United Nations. Scientific Committee on the Effects of Atomic Radiation. *Sources and effects of ionizing radiation: sources*, volume 1. United Nations Publications, 2000.

[58] G. F. Knoll. *Radiation detection and measurement*. John Wiley & Sons, 2010.

[59] J. H. Ely, R. T. Kouzes, B. D. Geelhood, J. E. Schweppe, and R. A. Warner. Discrimination of naturally occurring radioactive material in plastic scintillator material. In *Nuclear Science Symposium Conference Record, 2003 IEEE*, volume 2, pages 1453–1457. IEEE, 2003.

[60] N. Salikhov and O. Kryakunova. An increase of the soft gamma-ray background by precipitations. In *International Cosmic Ray Conference*, volume 11, page 367, 2011.

[61] R. J. Livesay, C. S. Blessinger, T. F. Guzzardo, and P. A. Hausladen. Rain-induced increase in background radiation detected by radiation portal monitors. *Journal of environmental radioactivity*, 137:137–141, 2014.

[62] S. Minato. Analysis of time variations in natural background gamma radiation flux density. *Journal of Nuclear Science and Technology*, 17(6):461–469, 1980.

[63] J. A. Hanley and B. J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1):29–36, 1982.

[64] C. E. Metz. Basic principles of ROC analysis. In *Seminars in nuclear medicine*, volume 8, pages 283–298. Elsevier, 1978.

[65] B. W. Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

[66] J. Aldrich. R.A. Fisher and the making of maximum likelihood 1912-1922. *Statistical science*, 12(3):162–176, 1997.

[67] S. S. Rao. *Engineering optimization: theory and practice*. John Wiley & Sons, 2009.

[68] H. Wan, T. Zhang, and Y. Zhu. Detection and localization of hidden radioactive sources with spatial statistical method. *Annals of Operations Research*, 192(1):87–104, 2012.

[69] B. Deb. Iterative estimation of location and trajectory of radioactive sources with a networked system of detectors. *IEEE Transactions on Nuclear Science*, 60(2):1315–1326, 2013.

[70] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT press, 2016. `http://www.deeplearningbook.org`.

[71] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Cognitive modeling*, 5(3):1, 1988.

[72] Y. LeCun. Generalization and network design strategies. In *Connectionism in perspective*, volume 19. Citeseer, 1989.

[73] Y. LeCun and Y. Bengio. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995, 1995.

[74] M. H. Jeong, C. J. Sullivan, M. Cheng, and S. Wang. Minimization of the impact of sensor velocity on the probability of source detection using geographically weighted methods. In *Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD), 2016*, pages 1–2. IEEE, 2016.

[75] M. Cheng. Analysis of a custom Android app designed to utilize cloud infrastructure in support of mobile sensor networks. Master's thesis, University of Illinois at Urbana-Champaign, 2017.

[76] D. J. Mitchell and L. T. Harding. GADRAS isotope ID user's manual for analysis of gamma-ray measurements and API for Linux and Android. *SAND2014-3933*, May 2014.

[77] M. W. Rawool-Sullivan, J. K. Mattingly, D. J. Mitchell, and J. D. Hutchinson. Combined MCNP/-GADRAS simulation of HPGe gamma spectra. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE*, pages 944–947. IEEE, 2012.

[78] C. Rhykerd, D. J. Mitchell, and J. K. Mattingly. Benchmarks for GADRAS performance validation. *Sandia National Laboratories, SAND*, 6550:2009, 2009.

[79] J. B. Stinnett. *Automated isotope identification algorithms for low-resolution gamma spectrometers*. PhD thesis, University of Illinois at Urbana-Champaign, 2016.

[80] W. R. Tobler. A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1):234–240, 1970.

[81] M. L. Stein. *Interpolation of spatial data: some theory for kriging*. Springer Science & Business Media, 2012.

[82] P. Monestiez, L. Dubroca, E. Bonnin, J.-P. Durbec, and C. Guinet. Geostatistical modelling of spatial distribution of Balaenoptera physalus in the Northwestern Mediterranean Sea from sparse count data and heterogeneous observation efforts. *Ecological Modelling*, 193(3):615–628, 2006.

[83] L. M. McShane, P. S. Albert, and M. A. Palmatier. A latent process regression model for spatially correlated count data. *Biometrics*, pages 698–706, 1997.

[84] E. Bellier, P. Monestiez, and C. Guinet. Geostatistical modelling of wildlife populations: a non-stationary hierarchical model for count data. In *geoENV VII–Geostatistics for Environmental Applications*, pages 1–12. Springer, 2010.

[85] M. Ali, P. Goovaerts, N. Nazia, Z. Haq, M. Yunus, and M. Emch. Application of Poisson kriging to the mapping of cholera and dysentery incidence in an endemic area of Bangladesh. *International Journal of Health Geographics*, 5(1):45, 2006.

[86] N. S. Asmarian, A. Ruzitalab, K. Amir, S. Masoud, and B. Mahaki. Area-to-area Poisson kriging analysis of mapping of county-level esophageal cancer incidence rates in Iran. *Asian Pacific Journal of Cancer Prevention*, 14(1):11–13, 2013.

[87] P. Goovaerts. Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. *International Journal of Health Geographics*, 4(1):31, 2005.

[88] A. Reinhart. An integrated system for gamma-ray spectral mapping and anomaly detection. Bachelor's thesis, University of Texas at Austin, 2013.

[89] G. Matheron. Principles of geostatistics. *Economic geology*, 58(8):1246–1266, 1963.

[90] E. Bai, K. Yosief, S. Dasgupta, and R. Mudumbai. The maximum likelihood estimate for radiation source localization: Initializing an iterative search. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 277–282. IEEE, 2014.

[91] G. Cordone, R. R. Brooks, S. Sen, N. Rao, C. Q. Wu, M. L. Berry, and K. M. Grieme. Improved multi-resolution method for MLE-based localization of radiation sources. In *Information Fusion (Fusion), 2017 20th International Conference on*, pages 1–8. IEEE, 2017.

[92] B. Schölkopf, A. Smola, and K.-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural computation*, 10(5):1299–1319, 1998.

[93] Y. Bengio. Deep learning of representations for unsupervised and transfer learning. In *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, pages 17–36, 2012.

[94] F. Pedregosa, G. Varoquaux, A. Gramfort, V Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.

[95] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pages 265–283, 2016.

[96] H. Paulheim and R. Meusel. A decomposition of the outlier detection problem into a set of supervised learning problems. *Machine Learning*, 100(2-3):509–531, 2015.

[97] J. R. Quinlan et al. Learning with continuous classes. In *5th Australian joint conference on artificial intelligence*, volume 92, pages 343–348. World Scientific, 1992.

[98] P.-N. Tan et al. *Introduction to data mining*. Pearson Education India, 2007.

[99] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander. LOF: identifying density-based local outliers. In *ACM sigmod record*, volume 29, pages 93–104. ACM, 2000.

[100] F. T. Liu, K. M. Ting, and Z.-H. Zhou. Isolation forest. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, pages 413–422. IEEE, 2008.

[101] F. T. Liu, K. M. Ting, and Z.-H. Zhou. Isolation-based anomaly detection. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 6(1):3, 2012.

[102] D. J. Mitchell, L. T. Harding, G. G. Thoreson, and S. M. Horne. GADRAS detector response function. *Sandia National Laboratories, SAND*, 19465:2014, 2014.

[103] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.

[104] S. M. Omohundro. *Five balltree construction algorithms*. International Computer Science Institute Berkeley, 1989.